



TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

TEEMU VALLINAHO
MEKAANISTEN HÄLYTINÄÄNIEN TUNNISTAMINEN
Diplomityö

Tarkastaja: professori Ireneusz De-
fee, laboratorioinsinööri Heimo Iha-
lainen
Tarkastaja ja aihe hyväksytty
Tieto- ja sähkötekniikan tiedekunta-
neuvoston kokouksessa 7. marras-
kuuta 2012

TIIVISTELMÄ

TAMPEREEN TEKNILLINEN YLIOPISTO

Signaalinkäsittelyn ja tietoliikennetekniikka koulutusohjelma

VALLINAHO, TEEMU: Mekaanisten hälytinäänien tunnistaminen

Diplomityö, 47 sivua, 2 liitesivua

joulukuu 2012

Pääaine: Multimedia

Tarkastaja: professori Ireneusz Defee, laboratorioinsinööri (ASE/TTY) Heimo Ihalainen

Avainsanat: akustiikka, digitaalisen signaalin prosessointi, audio, FFT, autoregressiivinen menetelmä

Tässä työssä tutkitaan mekaanisia ääniä, tarkemmin erilaisten hälyttimien kuten palohälyttimien, vesihälyttimien ja mukana kannettavien erikoishälyttimien hälytysäänien ominaisuuksia ja miten niitä voitaisiin tunnistaa. Tarkoituksena on näiden perusteella luoda järjestelmä, joka oppii hälyttimien ominaisuuksia ja tunnistaa niiden äänet. Näiden lisäksi tutkitaan miten äänet käyttäytyvät erilaisissa tiloissa ja miten erilaiset materiaalit vaimentavat hälyttimien ääniä.

Työ on jaettu kahteen osaan. Ensimmäinen osa on kirjallisuusselvitys, missä käydään läpi hälyttimen äänen analysoinnin kannalta tärkeät teoriat ja menetelmät. Tärkeimpinä alueina ovat äänen tehospektrin luominen ja siitä eri taajuuksien tunnistamista ja äkkinäisen muutoksen havaitseminen signaalissa eli toisin sanoen kovan äänen tunnistaminen. Näiden lisäksi tärkeä osa on hahmontunnistuksella ja miten sen eri menetelmiä voitaisiin käyttää hyväksi hälyttimien tunnistamiseen. Ensimmäiseen osioon kuului myös erilaisten standardien selvittäminen eri hälyttimistä. Näitä löytyi palovaroittimista, joille on hyvin tarkasti määritelty ominaisuudet.

Toinen osa on hälyttimien hälytysäänillä tehtyt testit. Ensin jokainen hälytin analysoidaan ja niiden äänen ominaisuudet selvitetään. Näistä selviää mitä ominaisuuksia kannattaa mahdollisesti käyttää äänen tunnistamiseen. Hälyttimiä testattiin myös erilaisissa tilanteissa kuten pihalla, toimistossa ja erittäin kaikuvassa tilassa. Näin saatiin selville miten ääni käyttäytyy ja onko mahdollista selvittää äänen ominaisuuksia kaikkialla. Havaittiin, että äänen tunnistaminen on hankalaa, mikäli kaiku pääsee sotkemaan ääntä liikaa. Testejä tehtiin myös selvittääkseen miten eri materiaalit vaikuttavat hälyttimiin. Testit osoittavat, että ääni vaimenee huomattavasti mutta mikään yksi selkeä taajuusalue ei katoa kokonaisuudessaan. Havaittiin myös, että signaali saattaa heiketä niin paljon, että kovan äänen tunnistamiseen soveltuvat algoritmit eivät välttämättä havaitse sitä. Vaimennustestit tehtiin valkoisella kohinalla ja sinisignaaleilla, jonka taajuus nousee koko ajan tasaisesti tiettyyn taajuuteen asti. Hälyttimien käyttäytymistä testattiin myös kylmässä tilassa. Havaittiin, että kova pakkanen saattaa vaikuttaa hälyttimien toimintaan ja se on otettava huomioon.

ABSTRACT

TAMPERE UNIVERSITY OF TECHNOLOGY

Master's Degree Programme in Signal Processing and Communications

VALLINAHÖ, TEEMU: Recognition of Mechanical Alarm Sounds

Master of Science Thesis, 47 pages, 2 Appendix pages

December 2012

Major: Multimedia

Examiner: Professor Ireneusz Defee, Laboratory Manager (ASE/TTY) Heimo Ihalainen

Keywords: acoustics, digital signal processing, audio, FFT, autoregressive method

This thesis studies mechanical sounds, more specific the sound of different alarms like fire detectors, water detectors and personal alarm devices. Object is to find the characteristics of the alarm and find a way to recognize those sounds. Idea is to create a system that can learn the sound of alarms and recognize them later. This study also includes sections where attenuation effect of different materials affects to characteristics of a certain alarms device.

This thesis is divided in two parts. First is literature survey about different theories and methods that are important for analysis of sounds. Most important parts of that include power spectral density and sudden change detection. Sudden change detection is basically just observing when the amplitude of the sound rises suddenly. Pattern recognition is also a part of this thesis. Methods from pattern recognition can be used to recognize the sounds. First part includes also report about two different standards that involves fire detectors. The characteristics of fire detectors are very closely defined in these standards.

Second part is the test and analysis of the sounds of alarm devices. These test showed what kind of characteristics each alarm device had and how they could be retrieved. Alarms were also tested in different situation; outside, in an office, in a very echoing area. These tests showed how sound acts in different situations. Echo could possibly make sound recognition very difficult. Attenuation was also tested. These tests showed how different materials attenuate sound. White noise and sinusoidal wave that gradually increases its frequency up to 8000 Hz were used. Tests showed that no material attenuates any single frequency, higher end of the frequency scale, i.e. close to 8000 Hz, was attenuated more than the lower part. Test included also putting alarms to freezer and testing how this affects them. Very low temperatures might have some effect but not always. It is still something that must be taken to consideration.

ALKUSANAT

Tämä on Tampereen teknillisessä yliopistossa tehty diplomityö. Työ on laadittu Mittaustekniikan laitoksella Biisafe Oy:lle. Työssä olen saanut asiantuntevia neuvoja Kimmo Konkarikoskelta ja käytännön ohjausta Heimo Ihalaiselta Tampereen teknilliseen yliopiston Systeemitekniikan laitokselta. Näiden lisäksi haluan kiittää Ireneusz Defeetä, multimedian professoria Tampereen teknillisestä yliopistosta tämän diplomityön tarkastuksesta. Biisafe Oy:n kanssa on selvitetty lähtökohdat tähän työhön, selvitetty tutkimuksen seuraava askel ja selvitetty järkevimmät ratkaisut itse järjestelmän tekoon. Haluan kiittää edellä mainittuja avustaan ja mahdollisuudesta tehdä tämä diplomityö. Lisäksi haluan kiittää suklaan keksijää enkä halua kiittää sitä, joka keksi kirjaston myöhästymismaksut.

SISÄLLYS

1	Johdanto.....	1
2	Hälyttimet ja standardit	2
2.1	Standardit	2
2.2	Hälyttimet	2
3	Digitaalinen signaalin analysointi.....	4
3.1	Digitaalinen signaali	4
3.2	Spektrin estimointi	6
3.2.1	Fourierin muunnos.....	6
3.2.2	Nopea Fourierin muunnos.....	7
3.2.3	Ikkunointi	7
3.2.4	Welchin menetelmä	11
3.3	Autoregressiivinen malli, AR	12
3.3.1	Ar- kertoimien laskeminen.....	14
3.3.2	Tehospektrin laskeminen	17
3.4	Signaalin tunnistaminen spektristä.....	18
3.4.1	Tehospektrien etäisyys.....	18
3.5	Suurimman energian alue	20
3.6	Signaalin tunnistaminen spektrogrammista	23
3.6.1	Spektrogrammi	23
3.7	Äkkinäiset muutokset signaalissa.....	25
3.8	Signaalin energia ja teho.....	25
3.8.1	Muutoksen havaitseminen energiasta ja tehosta	26
4	Hahmontunnistus	29
4.1	Hälyttimien äänten analysointi.....	31
4.2	Hälyttimet ja niiden ominaisuudet	35
5	Testaus.....	37
5.1	Hälyttimien testaus	37
5.2	Laitteiden aiheuttama vaimennus.....	37
5.3	Materiaalien aiheuttama vaimennus	38
5.4	Valkoinen kohina	38
5.5	Käden vaikutus.....	39
5.6	Heittotesti.....	39
5.7	Piha	39
5.8	Toimisto	40
5.9	Erittäin kaikuva tila	40
5.10	Hälytinlaitteiden testaus.....	41
6	Yhteenveto.....	43
	Lähteet.....	45
	Liite 1 Tehospektrien COSH- ja Itakura-Saito- algoritmilla lasketut etäisyydet	

TERMIT JA NIIDEN MÄÄRITELMÄT

AR	Autoregressiivinen malli. Voidaan kuvata ja ennustaa signaaleita.
dB(A)	Desibeli on äänen voimakkuuden mitta. Suluissa oleva A kertoo mitä painotussysteemiä on käytetty.
DFT	Discrete Fourier transform, diskreetti Fourierin muunnos.
FFT	fast Fourier transform, nopea Fourier-muunnos.
Hz	hertsi. Taajuuden yksikkö, $1 \text{ Hz} = 1/\text{s}$ eli yksi tapahtuma sekunnissa
LPC	Linear predictive coding
LSM	Least squares method, pienimmän neliösumman menetelmä
ms	Millisekunti, 0.001 sekuntia
Pa	Pascal, paineen yksikkö, $1 \text{ Pa} = 1 \text{ N/m}^2 = 10^{-5} \text{ bar}$
PSD	Power Spectral Density, kts. tehospektri
RMS	Root mean square, neliöllinen keskiarvo.
s	sekunti
STFT	short-time Fourier transform, lyhytaikainen nopea Fourierin muunnos
Tehospektri	Kuvaa signaalin tehon jakautumista eri taajuuskaistoille.

1 JOHDANTO

Tämän diplomityön tarkoitus on selvittää mekaanisten äänien ominaisuuksia, miten niitä voidaan havaita ja miten erilaiset mekaaniset äänet voidaan tunnistaa. Mekaaniset äänet ovat suhteellisen yksinkertaisia ja niiden ominaisuudet ovat toistuvia. Työn tarkoituksena on myös tarkkailla miten erilaiset äänet vaimenevat ja miten niiden ominaisuudet muuttuvat mikäli niitä on vaimennettu jollain tavoin. Mekaanisia ääniä tutkitaan nauhoittamalla niitä digitaaliseen muotoon ja eri menetelmillä tutkimalla niitä. Työssä kaikki signaalit ovat digitaalisia ja analogisia signaaleita ei ole kuin korkeintaan hyvin viitteellisellä tasolla. Tutkimuksissa haetaan ominaisuuksia sekä silmämääräisesti erilaisista kuvaajista että algoritmeilla laskemalla. Tutkimuksissa myös nauhoitetaan eri materiaaleilla nauhoitettuja mekaanisia ääniä. Tärkein työkalu on spektrogrammi, joka lasketaan jokaiselle eri äänelle. Spektrogrammista näkee hyvin äänen ominaisuudet ja miten materiaali vaikuttaa siihen. Materiaalit saattavat vaimentaa eri taajuuksia eri tavoin ja tämän vuoksi tehdään testejä myös valkoisella kohinalla ja sinisignaaleilla, jonka taajuus kohoa tasaisesti tietyn ajan. Kaikki algoritmit toteutetaan Matlab-ohjelmalla, samaten kaikki kuvaajat ovat tehty Matlab-ohjelman omilla rutiineilla.

Kappaleessa 2 käydään läpi erilaiset standardit, jotka liittyvät palovaroittimiin ja kerrataan millaisia vaatimuksia hälyttimen äänen suhteen niillä on. Tämän lisäksi kappaleessa käydään läpi nopeasti millaisia hälyttimiä voi olla ja mihin niitä mahdollisesti voidaan käyttää. Kappaleessa 3 kerrataan mikä on digitaalinen signaali ja miten sen eri ominaisuudet voidaan selvittää. Tähän liittyy nopea Fourierin muunnos, tehospektri, autoregressiivinen menetelmä ja spektrogrammi. Näiden lisäksi käydään läpi miten signaalin voidaan tunnistaa näiden avulla. Kappaleessa 3 käsitellään myös AR-kertoimia, signaalin energiasta ja tehosta sekä äkkinäisen muutoksen havaitsemista signaalista. Äkkinäisellä muutoksella tarkoitetaan tässä tapauksessa signaalin tehon tai energian muutosta eli periaatteessa äänisignaalin kohdalla, äänen voimakkuuden äkkinäistä kasvua. Kappaleessa 4 on hahmontunnistuksen perusteita. Signaalia tunnistaminen eri ominaisuuksista on nimenomaan hahmontunnistuksellinen ongelma. Kappaleessa 5 on erilaisten testien tulokset, mitä tätä diplomityötä varten tehtiin. Kaikkia tuloksia ei voida julkaista.

2 HÄLYTTIMET JA STANDARDIT

Tässä kappaleessa kerrotaan erilaisista hälyttimistä ja niihin liittyvistä standardeista. Hälyttimistä ja standardeista käsitellään vain ne, jotka ovat tämän diplomityön ja siihen liittyvän järjestelmän kannalta oleellisia.

2.1 Standardit

Suomessa on kaksi hälyttimiä koskevaa standardia; SFS-EN-14604 ja SFS-EN-54. (SFS-EN-14604). SFS-EN-14604 määrittelee testimenetelmät, suoritusarvot ja valmistajien ohjeiden vaatimukset koti- tai vastaavaan asuinkäyttöön tarkoitettuihin palovaroittimiin, jotka toimivat hajavalo-, valonläpäisy- tai ionisaatioperiaatteella. Standardissa määritellään testausta varten seuraavat ympäristöolosuhteet: lämpötila 15 °C...35 °C, suhteellinen kosteus 25 %...75 % ja ilmanpaine 86 kPa...106 kPa. Näiden perusteella voidaan luoda oletamus, että palovaroittimien on toimittava kyseisissä ympäristöolosuhteissa normaalilla tavalla ja joka kerta samalla tavalla. Samaten standardissa on määriteltä muita ympäristöön liittyviä testejä kuten värinäntesti ja herkkyys palosuureille. Nämä kaikki testit takaavat sen, että palovaroittimet ovat lähtökohtaisesti toimivia ja ne toimivat tiettyjen vaatimusten mukaisesti. Tämän diplomityön kannalta tärkeimmät tiedot ovat hälytysäänen äänenvoimakkuuden ja hälytysäänen taajuuden vaatimukset. Äänenvoimakkuuden tulee olla pienimmilläänkin 82 dB(A) kolmen metrin päässä hälyttimestä neljän minuutin jälkeen ja pienin nimellistaajuus ei saa ylittää 3500 Hz. (SFS-EN 14604 2006)

Standardi SFS-EN-54 määrittelee palohavaintsimien ja palohälytinjärjestelmän toiminnan [SFS-EN 54 2011]. Kyseessä on ensisijaisesti järjestelmät, jotka saavat käyttötehonsa ulkoisesta lähteestä, eikä sisäisestä paristosta tai akusta kuten standardin SFS-EN-14604 mukaiset hälyttimet. Tässä standardissa ei mainita erikseen taajuusalueita, mutta valmistajan on ilmoitettava kyseiset tiedot (SFS-EN 54 2006).

2.2 Hälyttimet

Hälyttimiä on luotu useisiin eri käyttötarkoituksiin. Niiden toimintaperiaate, hälytysääni ja voimakkuus vaihtelevat. Kaikista ei ole luotu standardia, niin kuin palovaroittimista on luotu, joten hälytinten varoitusäänet saattavat vaihdella hyvin paljon. Jokainen tämän työn puitteissa huomioitu hälytin tuottaa pääasiallisena varoituksena äänimerkin. Näitä äänisignaaleita tarkkailemalla, voidaan tunnistaa hälyttävä hälytin.

Palovaroitin: Tarkoitetaan nimenomaan sisäisellä voimanlähteellä eli esimerkiksi paristolla toimivat palovaroittimet. Tämä on hyvin yleinen varoitin, joka löytyy useista kohteista. Standardi SFS-EN-14604 määrittää näille varoittimille vaatimuksia. Varoitusaäni on hyvin selkeästi rajoittunut yhdelle taajuudelle ja siksi näiden tunnistaminen on suhteellisen helppoa. Hälyttimen taajuus on standardissa määrätty alle 3500 Hz. Hälyttimen ääni ei ole kokoaikainen vaan se vaikenee hyvin lyhyeksi hetkeksi aiheuttaen piipittävän äänen. (SFS-EN 14604 2006)

Vesihälytin: Hälytin, joka laukeaa, kun sen pohjassa olevat anturit johtavat sähköä eli ne yhdistyvät esimerkiksi veden välityksellä.

Kaasuhälytin: Reagoi ilmassa olevaan kaasuun ja antaa äänisignaalin, kun kaasua havaitaan. Kaasuhälyttimissä voi olla hyvin samankaltainen ääni kuin palovaroittimissa eli ääni voi olla voimakas yhdellä kapealla taajuuskaistalla.

Erikoishälyttimet: Erikoishälyttimet ovat yleensä mukana kannettavia pieniä hälyttimiä, jotka antavat äänimerkin kun hälytin laukaistaan. Poikkeuksen näissä edellisiin on, että hälytin pitää manuaalisesti laukaista. Erikoishälyttimissä äänisignaali voi olla melkein millainen vain, sitä ei ole rajoitettu mitenkään. Tällaisen äänen tunnistaminen on hankalampaa kuin esimerkiksi palovaroittimen sillä yhtä selkeää taajuutta ei välttämättä löydy.

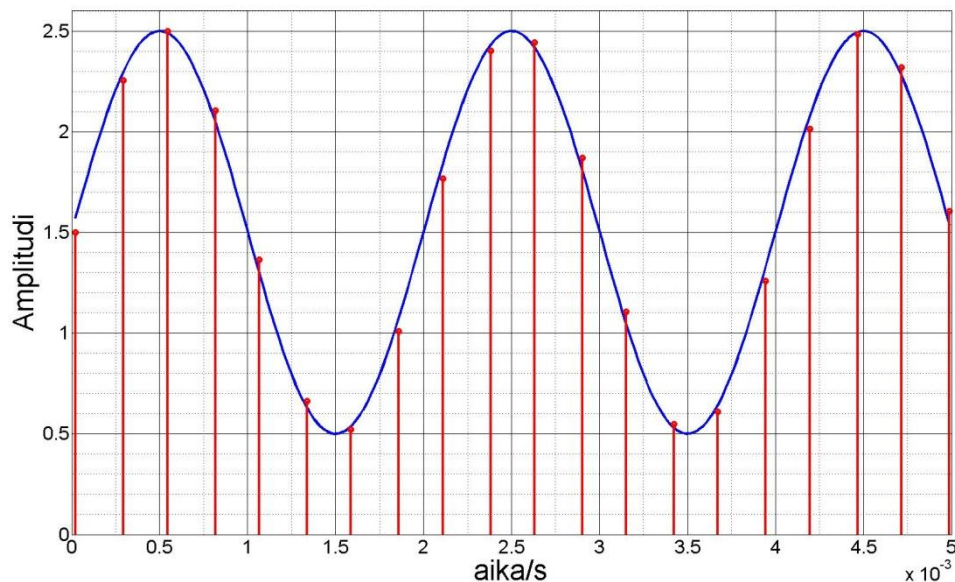
Näiden lisäksi on olemassa useita erilaisia hälyttimiä, joiden ääni vaihtelee hyvin runsaasti ja joiden tunnistamiseen pelkästään taajuuden perusteella ei ole mahdollista.

3 DIGITAALINEN SIGNAALIN ANALYSOINTI

Tässä kappaleessa kerrotaan digitaalisesta signaalista ja signaalin tunnistamisen eri keinoista. Keinot ovat rajattu sellaisiin, joita tässä työssä. Tässä työssä keskitytään pelkästään digitaalisiin signaaleihin, koska sovellusalue on hyvin voimakkaasti keskittynyt digitaalisen signaalin käsittelyyn ja käytettävissä ei ole analogista signaalia. Lisäksi signaalinkäsittely on digitaalisesti helpompaa kuin analogisesti.

3.1 Digitaalinen signaali

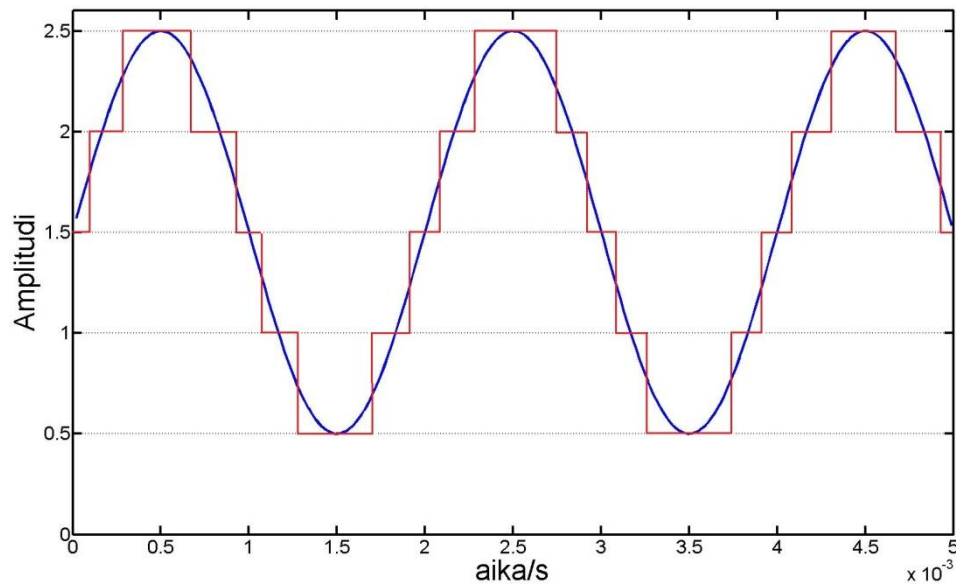
Digitaalisesta signaalista puhuttaessa on syytä ensin määritellä diskreetti-aikaisen signaalin ja jatkuva-aikaisen signaalin ero. Jatkuva-aikainen signaali voi saada tietyllä aikavälillä rajattoman määrän arvoja kun diskreetti-aikainen signaalille otetaan arvoja tietyin aikavälein. Tällöin puhutaan näytteistyksestä. Analogisesta signaalista otetaan näyte tietyin aikavälein. Tähän liittyy olennaisesti näytteistystaajuus eli näytteenottotaajuus (sampling rate) eli kuinka monta näytettä otetaan sekunnissa. Kuvassa 1 on jatkuva-aikainen signaali sinisellä ja punaisella on siitä otetut näytteen paikat tietyllä näytevälillä. (Oppenheim, et. al. 1975).



Kuva 1. Analoginen signaali, jossa näkyy aikaväli, joilla otetaan näyte signaalista.

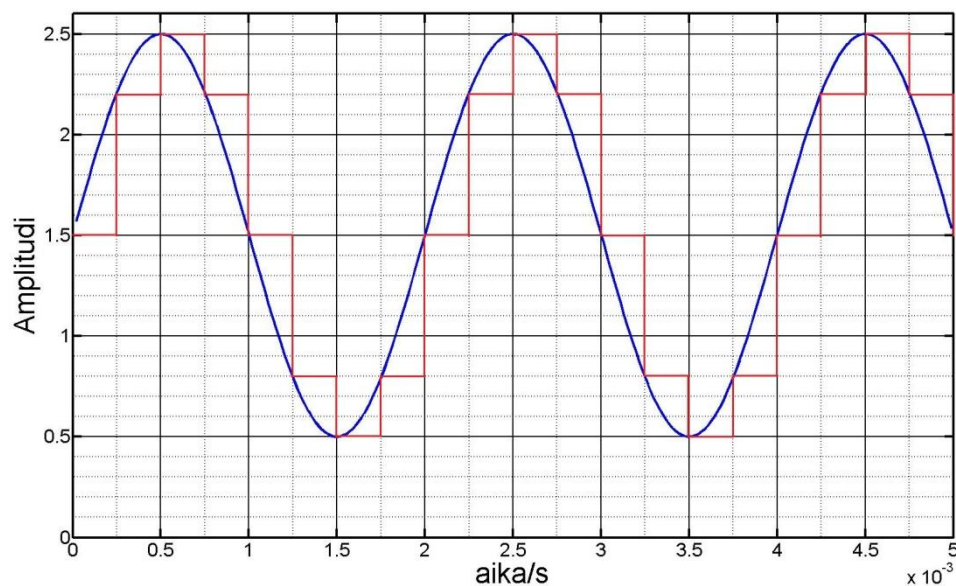
Digitaalinen signaali muodostetaan analogisesta signaalista ottamalla siitä näytteitä kuvan 1 punaisten nuolten kohdilta. Analoginen signaali voi saada rajattoman määrän arvoja. Digitaalinen signaali voi saada vain tiettyjä ennalta määrättyjä arvoja. Diskreetin

aikavälein otetut näytteet kvantisoidaan eli niille määrätään jonkin näistä ennalta määrätyistä arvoista sen mukaan mikä on lähimpänä todellista arvoa. Kuvassa 2 näkyy miten analogisen signaalin arvot voidaan kvantisoida. Sinisellä näkyy analoginen signaali ja punaisella on merkattu sen kunkin hetken arvo. Huomataan, että kvantisoitu arvo muuttuu portaittain ja voi saada vain tiettyjä arvoja.



Kuva 2. Analogisen signaalin kvantisointi.

Kuvassa 2 on siis analoginen signaali, joka saa jonkin arvon kaikilla ajan hetkillä. Signaali etenee hyppäyksin aina kahden kvantisointiarvon välillä. Lopullinen digitaalinen signaali muodostetaan ottamalla diskreetein aikavälein kvantisoitu analogisen signaalin arvo. Kuvassa 3 on esitetty miten digitaalinen signaali muodostuu. (Bellanger 1989).



Kuva 3. Diskreettiaikainen kvantisoitu signaali eli digitaalinen signaali punaisella.

Kuvasta 3 ei ole vaikea kuvitella, että digitaalinen signaali saa tietyin välein jonkin arvon eli esimerkiksi äänisignaalista muodostuu $1 \times n$ -kokoinen matriisi eli periaatteessa vain pitkä jono numeroita jotka tulkitaan äänisignaaleiksi. Sama pätee myös digitaaliseen kuvaan, joka on $m \times n$ - kokoinen matriisi.

3.2 Spektrin estimointi

Signaalin esitys aikatasossa ei useinkaan ole kovin hyödyllinen ja siitä on usein hankalaa irrottaa tietoa signaalin taajuuksista. Mielekkäämpää on käyttää signaalin spektriä, mistä selviää mistä taajuuksista signaali koostuu ja kuinka paljon kunkin taajuuden amplitudi on. (Huttunen 2005). Tehospektristä puhuttaessa tarkoitetaan signaalin tehon jakautumista eri taajuusalueille, siinä taajuusalue voi olla isompi kuin esimerkiksi 1 Hz. Tällaisesta esimerkki on Welchin menetelmä. Tässä työssä keskitytään erityisesti diskreetti aikaisten digitaalisignaalin kannalta oleellisiin asioihin ja jätetään jatkuva-aikaiset analogiset signaalit vähemmälle huomiolle.

3.2.1 Fourierin muunnos

Fouriern muunnos perustuu Jean Baptiste Joseph Fourierin päätelmään, että mikä tahansa jaksollinen signaali voidaan esittää sinimuotoisten signaalien summana. Tästä voidaan muodostaa Fourierin sarja (Smith, 1997). Fourierin muunnos voidaan tehdä mille vain signaalille, sillä voidaan ajatella, että jaksottomat signaalit ovat jaksollisia jos niiden jaksonaika on äärimmäisen pitkä. Tällöin Fourierin sarjan summat korvautuvat jatkuva-aikaisten signaalien kohdalla integraaleilla. (Oppenheim et. al. 1997). Fourierin muunnos voidaan esittää siis kaavana (Oppenheim et. al. 1997)

$$X(j\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt \quad (1)$$

ja kun käytämme Eulerin kaavaa

$$e^{jx} = \cos x + j \sin x \quad (2)$$

ja tietoa

$$\omega = 2\pi f \quad (3)$$

missä f on taajuus, saamme kaavan muotoon

$$X(j\omega) = \int_{-\infty}^{\infty} x(t) * (\cos 2\pi ft - j \sin 2\pi ft) dt \quad (4)$$

Kaavan (4) itseisarvon neliön odotusarvo on signaalin tehospektri eli

$$P_x(fj\omega) = E\{|X(j\omega)|^2\} \quad (5)$$

Tämän työn kannalta kiinnostavampi ja käytännöllisempi Fourier muunnos on diskreetti Fourier muunnos (DFT). Diskreetissä tapauksessa integraalit korvataan summilla ja näin ollen jatkuva-aikaisen signaalin Fourierin muunnos saadaan muotoon (Oppenheim et. al. 1997)

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n} \quad (6)$$

Tästä nähdään jokainen taajuuskomponentti mitä löytyy signaalista $x[n]$.

3.2.2 Nopea Fourierin muunnos

Fourierin muunnoksen laskeminen suoraan kaavasta (6) on hyvin hidas ja työläs prosessi. Mikäli signaalissa on vain yhtä taajuutta tai se on hyvin lyhyt, niin muunnoksen laskemien suoraan kaavasta voi olla mielekästä. Reaalimaailman signaalit eivät kuitenkaan ole yhden taajuuden signaaleita vaan niissä voi olla useita eri taajuuksia ja mukana saat-
taa olla kohinaa, joka vaikeuttaa entisestään laskutehtävää. Nopea Fourierin muunnos, FFT, keksittiin 1960-luvulla. FFT on nopea ja helppo tapa laskea Fourierin muunnos signaaleille. FFT osoittautui erittäin hyväksi työkaluksi diskreeteille digitaalisille signaaleille ja sitä käytetään laskemaan DFT lähestulkoon kaikissa käytännön sovelluksissa. Nimenomaan FFT:n ominaisuus laskea DFT nopeasti ja tehokkaasti on syy sen suosioon. (Oppenheim, et. al., 1997). Alun perin tämä algoritmi oli erityisen hyvä ja nopea silloin kun signaalin pituus, N , on kahden potenssi. (Bellanger 1989, s. 41) Tästä syystä signaalien pituus on hyvä olla kahden potenssina. Nykyään riittää, että signaalin pituus on alkulukujen tulo eli esimerkiksi $2^{N_1}3^{N_2}5^{N_3}$ ja $N_1+N_2+N_3 = N$.

3.2.3 Ikkunointi

Ikkunointi on signaalinkäsittelyssä käytetty menetelmä, missä signaali, $s(n)$, kerrotaan ikkunafunktiolla $w(n)$. Yleensä signaali $s(n)$ kerrotaan ikkunan ulkopuolella nolllalla ja

sen sisäpuolella jollain nollasta poikkeavalla arvolla (Aumala et. al. 1995). Ikkunoinnissa siis otetaan signaalin arvot tietyltä väliltä $-M \rightarrow M$ eli

$$p(n) = \begin{cases} x(n), & -M \leq n \leq M \\ 0, & \text{muulloin} \end{cases} \quad (7)$$

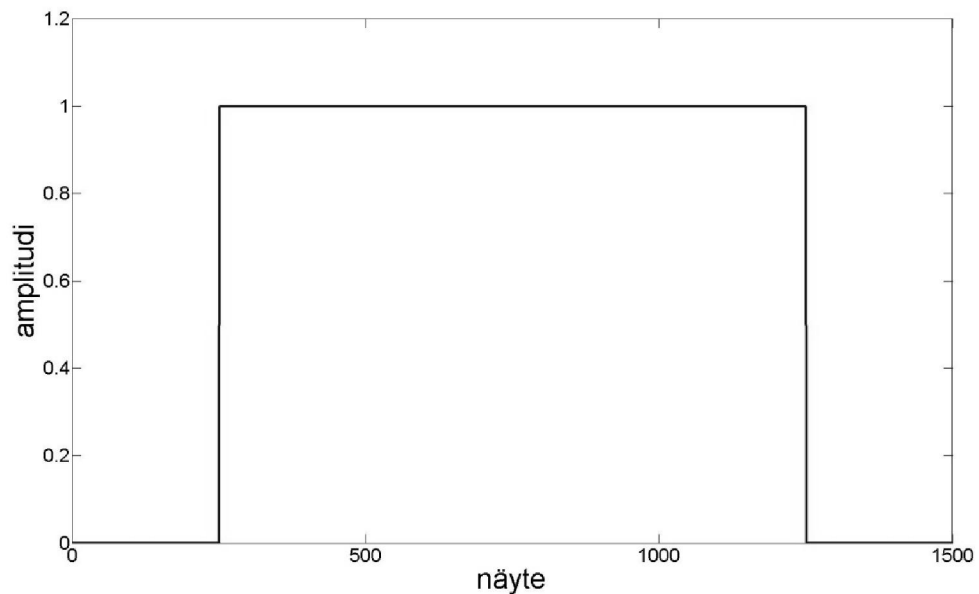
missä $-M \leq n \leq M$ on aikaikkuna (Oppenheim, et. al. 1997). Tämän jälkeen signaali kerrotaan ikkunafunktiolla $w(n)$ eli

$$x(n) = s(n)w(n) \quad (8)$$

missä $x(n)$ on uusi signaali. (Aumala et. al. 1995). Tällä tavoin signaalista saadaan sopivan mittainen osa analysointia varten. Ikkunointi-funktioita on useita erilaisia ja niillä on kaikilla omat hyvät ja huonot puolensa (Aumala et. al. 1995). Yksinkertaisin ikkunointifunktio on suorakulmainen ikkuna, missä ikkunan ulkopuolella olevat signaalin arvot kerrotaan nolllalla ja sisäpuolella kerrotaan ykkösellä eli signaali otetaan sellaisenaan käyttöön analysointia varten (Lynn et. al. 1994). Eli suorakaiteen muotoinen ikkuna on siis muotoa

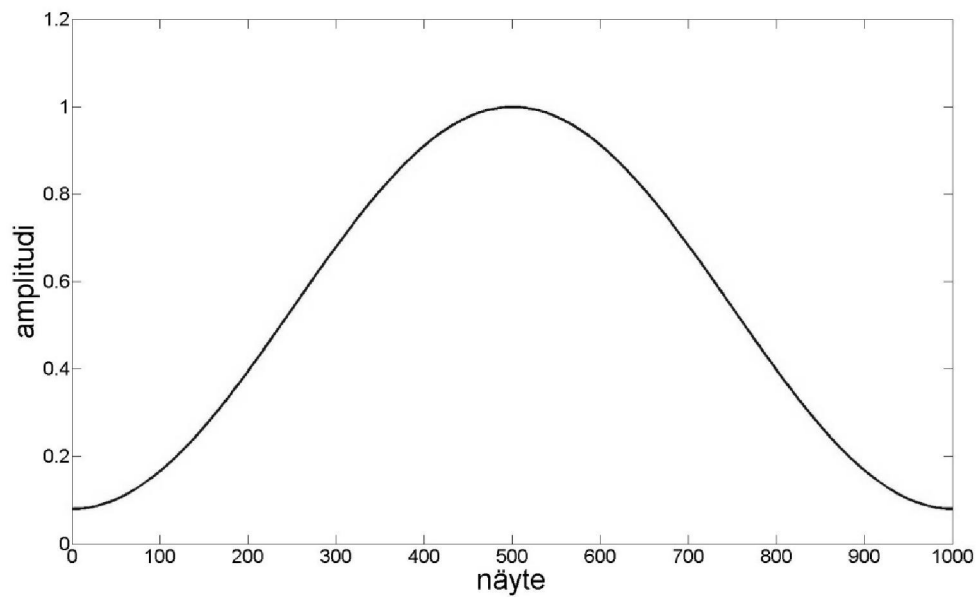
$$w[k] = \begin{cases} 1 & |k| \leq M \\ 0 & |k| > M \end{cases} \quad (9)$$

eli signaali tulee käyttöön sellaisenaan (Aumala et. al. 1994). Kuvassa 4 on visualisointi tällaisesta suorakaiteen muotoisesta ikkunasta.



Kuva 4. Suorakaiteen muotoinen ikkuna. Ikkunan rajalla tulee selkeä epäjatkuvuuskohta, mikä saattaa aiheuttaa ongelmia Fourierin muunnosta laskettaessa.

Yleensä signaali kuitenkin jatkuu ikkunan ulkopuolella ja suorakaiteen muotoisen ikkunan terävä reuna, missä signaalin arvot tipahtavat noltaan yllättäen ei ole paras mahdollinen. Tällaisella epäjatkuvuuskohdalla on ei-toivottuja vaikutuksia esimerkiksi taajuusspektrin laskemiseen (Lynn et. al. 1994). Tästä syystä ikkunaa saatetaan pehmentää hiukan vaihtelemalla ikkunafunktion kertoimia siten, että epäjatkuvuuskohta ei ole niin suuri kuin suorakaiteen muotoisessa ikkunassa (Lynn et. al. 1994) Tällaisia ikkunoita on esimerkiksi Hamming- ikkuna ja Hanning- ikkuna (Kay 1988). Kuvassa 5 on esitetty Hamming- ikkuna.



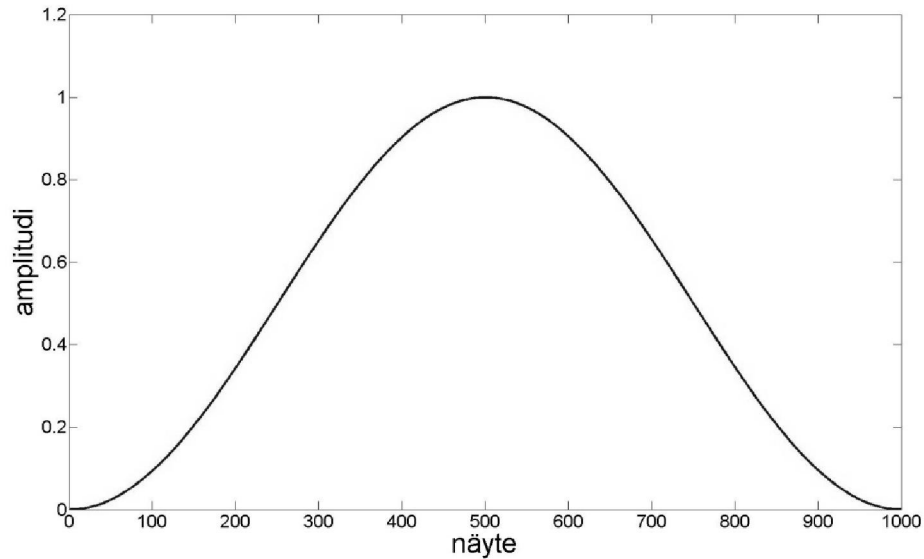
Kuva 5. Hamming- ikkuna. Ikkunan kertoimia on pehmennetty ja signaalin arvoja muokataan hiukan.

Hamming- ikkuna voidaan esittää muodossa

$$w[k] = \begin{cases} 0.54 + 0.46 \cos \frac{\pi k}{M} & |k| \leq M \\ 0 & |k| > M \end{cases} \quad (10)$$

jossa, M on haluttu aikaikkuna (Kay 1988).

Kuvassa 6 on esitetty Hanning- ikkuna.



Kuva 6. Hanning-ikkuna. Ikkunan arvot saavuttavat nollan aivan ikkunan rajalla toisin kuin esimerkiksi Hamming-ikkunassa.

Hanning- ikkuna voidaan esittää, samalla lailla kuin Hamming- ikkuna, muodossa

$$w[k] = \begin{cases} \frac{1}{2} + \frac{1}{2} \cos \frac{\pi k}{M} & |k| \leq M \\ 0 & |k| > M \end{cases} \quad (11)$$

(Kay 1988). Huomataan, että ero Hamming- ja Hanning- ikkunoiden välillä ei ole suuren suuri. Kuvista nähdään myös, että Hamming- ja Hanning- ikkunoiden kertoimet lähestyvät nollaa ikkunan reunoilla ja näin ollen reunalla oleva epäjatkuvuuskohta on huomattavasti pienempi kuin esimerkiksi suorakaiteen muotoisessa ikkunassa. Joissakin menetelmissä on syytä suorittaa signaalin jako datasegmentteihin siten, että segmentit ovat osakseen päällekkäin. Kuvista 5 ja 6 nähdään, että Hamming- ja Hanning- ikkunoilla painotetaan datasegmentin keskiosaa enemmän kuin reunoja. Tätä tilannetta helpotetaan sijoittamalla segmentit osaksi päällekkäin. Huomioitavaa on, että suorakaiteen muotoisessa ikkunassa ei päällekkäisyys ole tarpeen. (Aumala et. al. 1995). Kuvat 4,5 ja 6 ovat luotu Matlabin omilla rutiineilla.

3.2.4 Welchin menetelmä

Welchin menetelmällä voidaan estimoida signaalin tehoa eri taajuuksilla eli tehospektri, PSD, tähän käytetään hyödyksi kappaleessa 3.2.1 esitettyä Fourierin muunnosta (Oppenheim et al., 1975). Welchin menetelmässä saadaan poistettua kohinaa tehospektristä mutta toisaalta taajuusresoluutio pienenee (Welch 1970). Welchin menetelmän hyötynä on tarvittavien laskujen pienempi määrä, mikä vähentää tarvittavaa laskentatehoa ja ennen kaikkea laskenta-aikaa, sekä tarvittavan tallennustilan määrää. Welchin menetelmässä taajuudella olevan tehon varianssin kasvaessa, sen taajuusresoluutio tarkentuu. Tämä toimii myös toisinpäin. Eli mitä tarkempi taajuusresoluutio on, sitä enemmän taajuudella olevan tehon arvo voi vaihdella eli sen varianssi kasvaa. Welchin menetelmällä näitä on helppo muokata muuttamalla näytteiden määrää. (Welch 1975). Welchin menetelmä on muokattu aikaisemmin olemassa olleen Bartletin menetelmän pohjalta (Oppenheim et al., 1975).

Welchin menetelmässä signaali jaetaan K :hon osaan, joissa kaikissa on M näytettä eli

$$x^{(i)}(n) = x(n + iM - M), \quad 0 \leq n \leq M - 1, 1 \leq i \leq K \quad (12)$$

(Oppenheim, et al. 1975). Huomioitavaa on, että segmenttien pituus, M , määrittelee tehospektrin taajuusresoluution

$$\Delta f = \frac{1}{T} = \frac{f_s}{M} \quad (13)$$

(Aumala et. al. 1995.).

Tässä vaiheessa signaalista saatuun datasegmenttiin käytetään ikkunointi-funktiota $w(n)$. Tämän jälkeen lasketaan periodogrammi eli toisin sanoen taajuusspektri. Tämä voidaan merkitä:

$$J_M^{(i)}(\omega) = \frac{1}{MU} \left| \sum_{n=0}^{M-1} x^{(i)}(n) w(n) e^{-j\omega n} \right|^2, \quad i = 1, 2, \dots, K \quad (14)$$

(Oppenheim, et al. 1975). Tässä kaavassa:

$$U = \frac{1}{M} \sum_{n=0}^{M-1} w^2(n) \quad (15)$$

(Oppenheim, et al. 1975). Tehospektri on näiden periodogrammien keskiarvo (Welch 1975) joka lasketaan siis kaavalla

$$B_{xx}^w = \frac{1}{K} \sum_{i=1}^K J_M^{(i)}(\omega) \quad (16)$$

(Oppenheim et al. 1975). Taajuusarvot, jotka saadaan tähän tehospektrin estimaattiin mukaan, voidaan laskea kaavasta

$$f_k = \frac{k}{T} = \frac{k f_s}{M} \quad (17)$$

jossa $k = 0, 1, 2, \dots, M-1$, M on datasegmentin pituus ja f_s on signaalin näytteenottotaajuus (Aumala et. al. 1995).

Periodogrammien laskemiseen voidaan käyttää nopeaa Fourierin muunnosta (Kay 1988), mikä nopeuttaa laskutoimitusta huomattavasti. Tämä onkin nykyään käytössä hyvin laajasti. Welchin menetelmässä siis lasketaan signaalin pienien osien taajuusspektri ja lasketaan jokaiselle taajuusalueelle keskiarvo näistä kaikista pienistä osista. Welchin menetelmässä on hyväksi käyttää osittain päällekkäisiä ikkunoita (overlapping window), että kaikki signaalin arvot saadaan muuttumattomina mukaan tehospektriin (Kay 1988). Welchin menetelmä on hyvin laajasti teollisuudessa käytetty menetelmä juuri sen nopeuden takia. Yleensä ei ole tarpeen estimoida tehospektriä äärimmäisen tarkasti vaan pienellä taajuusresoluutiolla laskettu estimaatti on riittävä hyvin moneen sovellukseen. Welchin menetelmällä saatu tehospektri on siis jaettu lyhyisiin tasamittaisiin taajuuskaistoihin. Signaalin spektri, joka on jatkuva välillä nollasta 22050 Hz:iin, voisi olla Welchin menetelmän jälkeen jaettuna 129 osaan, tämä on aiemmin mainittu taajuusresoluutio. Tällöin yhden taajuuskaistan leveydeksi tulisi kaavaa (17) käyttäen noin 344Hz. Eli jokin taajuuskaista olisi $f \text{ Hz} \dots f+344\text{Hz}$ ja tälle laskettaisiin teho.

3.3 Autoregressiivinen malli, AR

Autoregressiivinen malli yrittää ennustaa signaalin tulevia arvoja sen edellisistä arvoista,

$$y(m) = \sum_{k=1}^p a_k y(m-k) + u(m) \quad (18)$$

jossa, $u(m)$ on signaalissa esiintyvä kohina ja a_k ovat reaalisia kertoimia ja p on haluttujen kertoimien määrä (Vaseghi 2000). Signaalin y seuraava arvo, $y(m)$, lasketaan siis kertomalla $y:n$ p edellistä arvoa kertoimella a_k ja laskemalla nämä arvot yhteen. AR-malli on näin ollen lineaarista ennustamista, joka tunnetaan digitaalisessa signaalinkäsittelyssä nimellä LPC (linear predictive coding). Jotta nähtäisiin AR-mallin tiettyjä ominaisuuksia, on syytä johtaa ensin kaavan z -muunnos. Aloitamme johtamalla signaalin yleisen siirtofunktion (Huttunen 2005)

$$y(n) = \sum_{k=0}^K b_k x(n-k) + \sum_{m=1}^M a_m y(n-m) \quad (19)$$

Ottamalla tästä z -muunnos saadaan

$$Y(z) = \sum_{k=0}^K b_k x(z) z^{-k} + \sum_{m=1}^M a_m Y(z) z^{-m} \quad (20)$$

Tätä voidaan hiukan muokata

$$Y(z) \left(1 - \sum_{m=1}^M a_m z^{-m} \right) = X(z) \sum_{k=0}^K b_k z^{-k} \quad (21)$$

Tästä voidaan johtaa ulostulon z -muunnos

$$Y(z) = \frac{\sum_{k=0}^K b_k z^{-k}}{1 - \sum_{m=1}^M a_m z^{-m}} X(z) \quad (22)$$

Siirtofunktio on yleisesti muotoa

$$H(z) = \frac{Y(z)}{X(z)} \quad (23)$$

eli siirtofunktio on siis

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{k=0}^K a_k z^{-k}}{1 - \sum_{m=1}^M b_m z^{-m}} \quad (24)$$

Kun muutamme kaavan (24) osoittajan arvoksi 1, saamme sen muotoon

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{1 - \sum_{m=1}^M b_m z^{-m}} \quad (25)$$

(Mitra 2006). Tämä on AR- mallin siirtofunktio. Siitä huomataan, että AR- malli on IIR- suodin (infinite impulse response) ja niin sanottu all-pole suodin (Mitra 2006). Ensimmäinen kerroin voidaan asettaa ykköseksi, $b_0 = 1$ (Porat 1994). AR- mallia voidaan käyttää useissa eri sovelluksissa, kuten spektri analyysissa, järjestelmän tunnistamisessa, puheanalyysissa, pakkauksessa ja suodinsuunnittelussa (Mitra 2006).

3.3.1 Ar- kertoimien laskeminen

AR- mallissa on tärkeää kuinka monta kerrointa sille lasketaan. Tämä taas riippuu siitä, mitä AR- mallilla on tarkoitus tehdä. Koska tämän diplomityö kannalta tärkeintä on tehospektrin laskeminen, otamme sen lähtökohdaksi.

Kertoimien määrä riippuu siitä millainen signaali halutaan ennustaa. Yksinkertainen yhden taajuuskomponentin sinisignaali on helpompi ennustaa kuin esimerkiksi monimutkainen useita taajuuskomponentteja sisältävä ihmisen puhesignaali. AR- kertoimien laskemiseen on olemassa useita eri keino kuten pienimmän neliösumman menetelmä, Yule-Walker- menetelmä ja Levinson-Durbin algoritmi. (Porat 1994). Käydään läpi Yule-Walker ja huomattavasti paremmin tämän diplomityön kannalta sopivampi pienimmän neliösumman menetelmä eli LSM.

Yule-Walker- menetelmä lähtee liikkeelle siitä, että y_t ennustettu arvo on

$$\hat{y} = - \sum_{k=1}^p a_k y_{t-k} \quad (26)$$

(Porat 1994). Yhtälössä 24 p on haluttujen kerrointen määrä. Kyseessä on siis y :n estimaatti \hat{y} . Tästä voimme muodostaa Yule-Walkerin yhtälöt

$$\begin{bmatrix} r_0 & \cdots & r_p \\ \vdots & \ddots & \vdots \\ r_p & \cdots & r_0 \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} \sigma_u^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (27)$$

missä $\{r_0, r_1, \dots, r_p\}$ ovat kovariansseja. Mikäli nämä ovat tiedossa, on yhtälön tuntemattomat eli $\{\sigma_u^2, a_1, \dots, a_p\}$. (Porat 1994) Nämä kovarianssit voidaan korvata niiden estimaateilla, jotka saadaan laskettua kaavasta (Porat 1994)

$$\hat{r}_k = \hat{r}_{-k} = \frac{1}{N} \sum_{t=0}^{N-k-1} y_k + ty^t \quad (28)$$

Näin ollen voidaan kertoimet laskea kaavalla

$$\begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \vdots \\ \hat{a}_p \end{bmatrix} = - \begin{bmatrix} \hat{r}_0 & \hat{r}_1 & \cdots & \hat{r}_{p-1} \\ \hat{r}_1 & \hat{r}_0 & \cdots & \hat{r}_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{r}_{p-1} & \hat{r}_{p-2} & \cdots & \hat{r}_0 \end{bmatrix}^{-1} \begin{bmatrix} \hat{r}_1 \\ \hat{r}_2 \\ \vdots \\ \hat{r}_p \end{bmatrix} \quad (29)$$

Tuntemattomista on enää laskematta varianssi ja se saadaan kaavalla

$$\hat{\sigma}_u^2 = \hat{r}_0 + \sum_{k=1}^p \hat{r}_k \hat{a}_k \quad (30)$$

Yule-Walkeria- menetelmää kutsutaan myös autokorrelaatio menetelmäksi. Kaavassa (29) autokorrelaatio on nimenomaan korrelaatiota signaalin itsensä kanssa.

Toinen menetelmä on pienimmän neliösumman menetelmä. Siinä tarkoituksena on minimoida mallinnusvirheen $u_{t,p}$ neliösumman eli

$$\min l(a) = \sum_{t=1}^T u_{t,p}^2 \quad (31)$$

(Phongsaphen 2006). Pienimmän neliösumman estimaatti muuttujalle θ yleisessä muodossa on

$$\hat{\theta} = [H(Y)^T H(Y)]^{-1} H(Y)^T \mathbf{h}(Y) \quad (32)$$

(Porat 1994). Tämä on nimenomaan painotettu lineaarinen pienimmän neliösumman estimaatti muuttujalle θ . AR- kerrointen muodostamisessa muokataan kaavaa (19) ensin hiukan

$$y_t = \sum_{k=1}^p a_k y_{t-k} = u_{t,p} \quad (33)$$

Toinen tarvittava kaava on

$$y_t = \sum_{k=1}^p a_k y_{t-k} = \tilde{u}_{t,p} \quad (34)$$

Näistä kahdesta kaavasta voimme johtaa kaksi eri yhtälöryhmää, ensin yhtälöstä 29

$$\begin{bmatrix} y_p \\ y_{p+1} \\ \vdots \\ y_{N-1} \end{bmatrix} + \begin{bmatrix} y_{p-1} & y_{p-2} & \cdots & y_0 \\ y_p & y_{p-1} & \cdots & y_1 \\ \vdots & \vdots & \ddots & \vdots \\ y_{N-2} & y_{N-3} & \cdots & y_{N-p-1} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} u_{p,p} \\ u_{p+1,p} \\ \vdots \\ u_{N-p-1,p} \end{bmatrix} \quad (35)$$

Jossa N on näytteiden määrä (Porat 1994). Seuraavaksi yhtälöstä 30

$$\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{N-p-1} \end{bmatrix} + \begin{bmatrix} y_1 & y_2 & \cdots & y_p \\ y_{2p} & y_3 & \cdots & y_{p+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_{N-p} & y_{N-p+1} & \cdots & y_{N-1} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} \tilde{u}_{0,p} \\ \tilde{u}_{1,p} \\ \vdots \\ \tilde{u}_{N-p-1,p} \end{bmatrix} \quad (36)$$

Tiivistettynä kaava (35) on muotoa

$$\mathbf{y}_f + \mathbf{Y}_f \mathbf{a} = \mathbf{e}_f \quad (37)$$

ja kaava (36) on

$$\mathbf{y}_b + \mathbf{Y}_b \mathbf{a} = \mathbf{e}_b \quad (38)$$

Näistä saadaan kaavan (32) mukainen muoto.

$$\hat{\mathbf{a}}_f = -(\mathbf{Y}_f^T \mathbf{Y}_f)^{-1} \mathbf{Y}_f^T \mathbf{y}_f \quad (39)$$

ja

$$\hat{\mathbf{a}}_b = -(\mathbf{Y}_b^T \mathbf{Y}_b)^{-1} \mathbf{Y}_b^T \mathbf{y}_f \quad (40)$$

Kaavoissa (39) ja (40) olevat estimaatit, $\hat{\mathbf{a}}_f$ ja $\hat{\mathbf{a}}_b$, ovat niin sanottuja eteenpäin pienimmän neliösumman estimaatti ($\hat{\mathbf{a}}_f$) ja taaksepäin pienimmän neliösumman estimaatti ($\hat{\mathbf{a}}_b$) (Porat 1994). Nämä voidaan vielä yhdistää muotoon

$$\hat{\mathbf{a}}_{fb} = -(\mathbf{Y}_f^T \mathbf{Y}_f + \mathbf{Y}_b^T \mathbf{Y}_b)^{-1} (\mathbf{Y}_f^T \mathbf{y}_f + \mathbf{Y}_b^T \mathbf{y}_b) \quad (41)$$

Pienimmän neliösumman menetelmä sopii paremmin tehospektrin estimointiin kuin esimerkiksi Yule-Walker. Kaavan (41) mukainen eteen-taaksepäin pienimmän neliösumman menetelmä on yleensä pidetty parempana mutta sen laskemiseen tarvitaan kaksi kertaa enemmän laskutoimituksia kuin muiden. (Porat 1994).

3.3.2 Tehospektrin laskeminen

Autoregressiivisistä kertoimista voi laskea signaalin tehospektrin autokorrelaation avulla (Vaseghi 2000). Autokorrelaatio kertoo kuinka paljon signaali korreloi itsensä kanssa tietyllä viiveellä. Sillä pystyy havaitsemaan muun muassa signaalin jaksollisuutta ja sitä käytetään laskemaan signaalin tehospektriä. (Aumala 1995). Tässäkin tapauksessa Fourierin muunnos on tarpeen, sillä tehospektri on signaalin korrelaatiofunktion Fourierin muunnos. Autokorrelaatiofunktio r_{xx} noudattaa kaavaa

$$r_{xx}(m) = E[x(m)x(m+k)] \quad (42)$$

(Vaseghi 2000) missä k on signaalin viivästys. Käytännössä autokorrelaatiofunktio estimoidaan N näytteen mittaisesta signaalista

$$\hat{r}_{xx}(m) = \frac{1}{N-|m|} \sum_{k=0}^{N-|m|-1} x(k)x(k+m), \quad k = 0, 1, \dots, N-1 \quad (43)$$

jossa, m on viive, joka lähestyy signaalin pituutta N . (Vaseghi 2000). Viiveen suuria arvoja voidaan tasoittaa käyttämällä kolmion muotoista ikkunaa

$$w(m) = \begin{cases} 1 - \frac{|m|}{N}, & |m| \leq N-1 \\ 0, & \text{muulloin} \end{cases} \quad (44)$$

Kaavan (8) mukaisesti ikkunoitaessa alkuperäinen signaali kerrotaan ikkunointifunktiolla eli tässä tapauksessa kaava (43) kerrotaan kaavalla (44) ja saadaan

$$\hat{r}_{xx}(m) = \frac{1}{N} \sum_{k=0}^{N-|m|-1} x(k)x(k+m), \quad k = 0, 1, \dots, N-1 \quad (45)$$

Energian spektri voidaan määritellä diskreetti aikaisille signaaleille seuraavalla tavalla

$$S_x(f) = E\{|X(f)|^2\} = E\left\{\left|\sum_{m=-\infty}^{\infty} x(m)e^{-j2\pi fm}\right|^2\right\} \quad (46)$$

eli kyseessä on itseisarvon neliö. Kuten sanottua, tehospektri on signaalin $x(m)$ autokorrelaatiofunktion Fourierin muunnos ja tämä saadaan kaavasta (46)

$$|X(f)|^2 = P_{xx}(f) = \sum_{m=-\infty}^{\infty} r_{xx}(m) e^{-j2\pi f m} \quad (47)$$

missä $r_{xx}(m)$ on signaalin $x(m)$ autokorrelaatiofunktio. Autoregressiivisen prosessin, jota kuvaa siis kaava (18), tehospektri saadaan kaavasta

$$P_{xx}^{AR}(f) = \frac{\sigma_u^2}{\left|1 - \sum_{k=1}^p a_k e^{-j2\pi f k}\right|^2} \quad (48)$$

Kaavan (47) ei kuitenkaan ole pakollista, sillä autokorrelaatiokertoimet on mahdollista laskea suoraan AR- kertoimista ja näin ollen on mahdollista käyttää kaavaa (46) (Brookes 2012).

3.4 Signaalin tunnistaminen spektristä

Signaalin tehospektriä voidaan käyttää signaalien erottamiseen toisistaan. Jokaisella äänellä on erilainen taajuusjakauma ja näin ollen eri taajuuksien spektriä tutkimalla voidaan suorittaa signaalien jakoa toisiinsa. Tässä käydään läpi tämän diplomityön kannalta keskeisimmät tunnistamismenetelmät, jotka pohjautuvat tehospektriin.

3.4.1 Tehospektrien etäisyys

Tehospektri eli PSD on sovelluksissa käytännössä $1 \times N$ kokoinen matriisi, missä jokainen jokaisen alkion indeksi viittaa tiettyyn taajuuteen ja jokainen alkion arvo signaalin tehoon sillä taajuudella. Siihen voidaan siis käyttää tavallisia matriisin laskusääntöjä ja sitä voidaan täysin kohdella matriisina.

Tehospektrin etäisyys toisesta tehospektristä voi olla hyvin tehokas työkalu signaalien tunnistamisessa ja luokittelussa sillä tehospektrien etäisyys toisistaan voidaan tulkita tarkoittamaan samankaltaisuutta. Mitä pienempi etäisyys on, sitä suurempi on samankaltaisuus (Wei 2000). Näiden etäisyyksien laskemiseen tarvittava tehospektri voidaan laskea joko suoraan AR- kertoimista tai se voidaan laskea jollain muulla tavoin saadusta tehospektristä. Näitä on käyty läpi kappaleessa 3.2.2 tai Welchin menetelmä kappaleessa 3.2.5. Signaalin tehospektrien etäisyyksiin on olemassa useita erilaisia algoritmeja ja tässä käsitellään niistä muutama tämän työn kannalta tärkein.

Tehospektrien etäisyydet voidaan esittää muodossa $d(x, y)$, missä x ja y ovat tehospekt-
rit $x = P(s_1)$ ja $y = P(s_2)$, jotka on saatu jollain tavalla estimoitua. Saatu tulos vaihtelee
jonkin verran sen mukaan, mitä algoritmia käytetään. Algoritmin valinta taas riippuu
siitä millaista signaalia käytetään. Gray ja Markel esittelevät artikkelissaan algoritmeja,
jotka ovat symmetrisiä eli $d(x, y) = d(y, x)$ esimerkiksi COSH- algoritmin (Gray 1976)
mutta kaikkien kohdalla näin ei ole, esimerkiksi Itakura-Saito- algoritmi ei ole symmet-
rinen (Iser 2008). Näin ollen tulos vaihtelee, sen mukaan kumpaa tehospekttriä verrataan
kumpaan (Wei 2000). Symmetrisyydestä seuraa tuloksen yksiselitteisyys, mikä taas
helpottaa sen käyttöä hahmontunnistuksen osana. Toisaalta esimerkiksi Itakura-Saito-
algoritmi havaitsee tehospektristä piikit paremmin kuin symmetriset menetelmät, mikä
voi olla käyttökelpoista tietyissä tilanteissa. (Wei 2000) (Gray 1976). Tehospektrien
etäisyyden voi laskea myös AR- kertoimien avulla. Edellä on esitetty miten AR- ker-
toimista saadaan laskettua tehospekttri ja tämän avulla voidaan laskea suoraan etäisyyk-
siä (Brookes 2012).

COSH- etäisyys on sopiva spektrien etäisyyden laskemiseen diskreetissä tapauksessa,
mikä tekee siitä sopivan tämän diplomityön kannalta (Wei 2000). COSH- etäisyys sopii
myös hyvin tilanteisiin, missä on odotettavissa suuria eroja tehospektrissä ja tämä sopii
taas erittäin hyvin puheen tunnistamiseen (Gray 1976). COSH- etäisyys voidaan laskea
kaavalla

$$E_{cosh} = \frac{1}{2N} \sum_{m=1}^N \left[\frac{P(s_1)}{P(s_2)} - \log \frac{P(s_1)}{P(s_2)} + \frac{P(s_2)}{P(s_1)} - \log \frac{P(s_2)}{P(s_1)} - 2 \right] \quad (49)$$

(Wei 2000). Tämä on diskreetissä muodossa esitetty. Liitteessä 1 on taulukko, missä on
laskettu COSH- etäisyyksiä eri tilanteissa eri hälyttimillä. Tästä näkee hyvin, millaisia
etäisyyksiä ja tuloksia tällä menetelmällä voi saada. Tähän on käytetty Mike Brookesin
kehittää toolboxia Matlabille (Brookes 2012). Näistä huomataan, että eri menetelmät
tuottavat hieman erilaisia tuloksia. Huomataan, että hälyttimen etäisyys itseensä on aina
0. Tästä seuraavaksi pienin etäisyys pitäisi olla vähintään saman hälyttimen handsfreellä
nauhoitettu näyte tai vähintään saman ryhmän sisällä. Taulukoista nähdään, että hälytin
lähin on tosiaan sen handsfreellä nauhoitettu pari. Erityisesti erikoishälytinten kohdalla
tämä algoritmi tuottaa pienimmän etäisyyden toisiin erikoishälyttimiin. Muiden hälytin-
ten kohdalla saattaa jollakin erikoishälyttimellä olla pienempi etäisyys kuin toisilla sa-
mankaltaisilla. Esimerkiksi palohälytin 1 ja palohälytin 2 etäisyys on hiukan suurempi
kuin palohälytin 1 ja erikoishälyttimen 6 etäisyys.

Itakura-Saito- etäisyys on asymmetrinen spektrietäisyyden laskemiseen käytetty algo-
ritmi. Algoritmi soveltuu paremmin spektrin huippujen löytämiseen kuin COSH- algo-
ritmi. (Gray 1976). Toisaalta Itakura-Saito- algoritmi ei löydä niin helposti spektrissä
esiintyviä matalia kohtia eli niin sanottuja laaksoja (Wei 2000). Itakura-Saito-

menetelmä on kuitenkin melko yleisesti käytetty spektrien etäisyyden laskemiseen ja otetaan sen vuoksi mukaan tähän diplomityöhön. Itakura-Saito- algoritmi määrittellään

$$d_{IS}(A, \hat{A}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\left[\frac{|\hat{A}(e^{j\Omega})|^2}{|A(e^{j\Omega})|^2} - \ln \frac{|\hat{A}(e^{j\Omega})|^2}{|A(e^{j\Omega})|^2} \right] - 1 \right) d\Omega \quad (50)$$

(Iser 2008), missä $\hat{A}(e^{j\Omega})$ ja $A(e^{j\Omega})$ ovat signaalin tehospektri. Tämä voidaan esittää myös diskreetissä muodossa

$$E_{I-S} = \frac{1}{2} \sum_{m=1}^N \left[\frac{P(s_1)}{P(s_2)} - \log \frac{P(s_1)}{P(s_2)} - 1 \right] \quad (51)$$

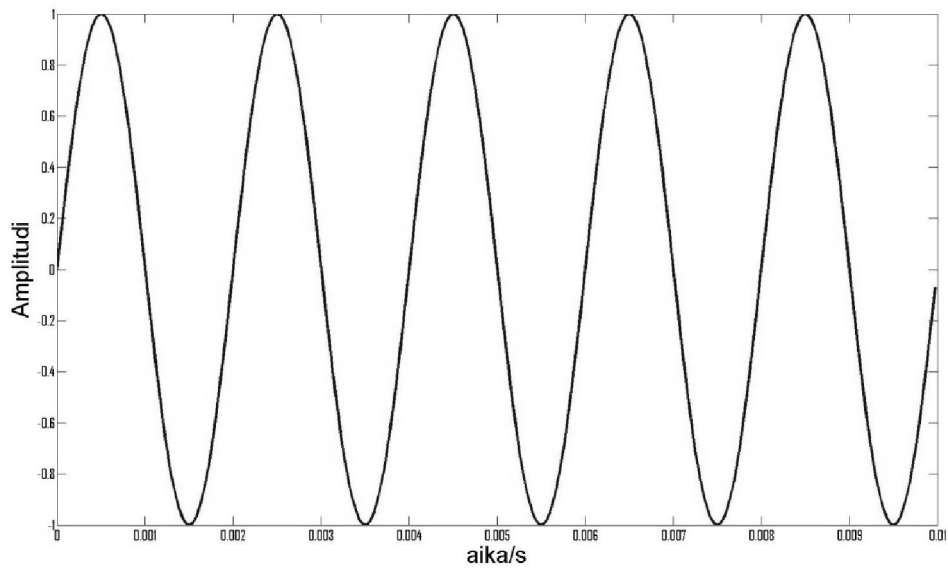
(Wei 2000), missä $P(s_1)$ ja $P(s_2)$ ovat signaalin tehospektri. Tästä huomataan ero COSH-algoritmiin. Molemmat hyödyntävät logaritmeja mutta Itakura-Saito laskee logaritmi ja etäisyyden vain toisin päin ja COSH laskee molemmiin päin. Tästä seuraa se, minkä vuoksi toinen on asymmetrinen ja toinen on symmetrinen. Liitteessä 1 sivulla 2 on taulukko Itakura-Saito tehospektrien etäisyyksistä eri hälyttimillä. Tehospektri on laskettu käyttämällä AR- kertoimia. Tulokset ovat hyvin samanlaisia kuin COSH- algoritmilla. Näiden kahden lisäksi on useita muitakin keinoja kuten Log Spectral Deviation, RMS-Log Spectral Deviation, Cepstral Distance ja Itakura. (Iser 2008).

3.5 Suurimman energian alue

Yksinkertaiset signaalit, joissa on vain yksi tai muutama komponentti, voidaan tunnistaa tehospektristä etsimällä taajuus, millä on suurin mahdollinen arvo. Tämä on toimiva silloin, kun signaalilla on yksi tai muutama selkeästi erottuva komponentti eli se muistuttaa hyvin selvästi sinimuotoista signaalia. Tällainen signaali voidaan merkitä muodossa

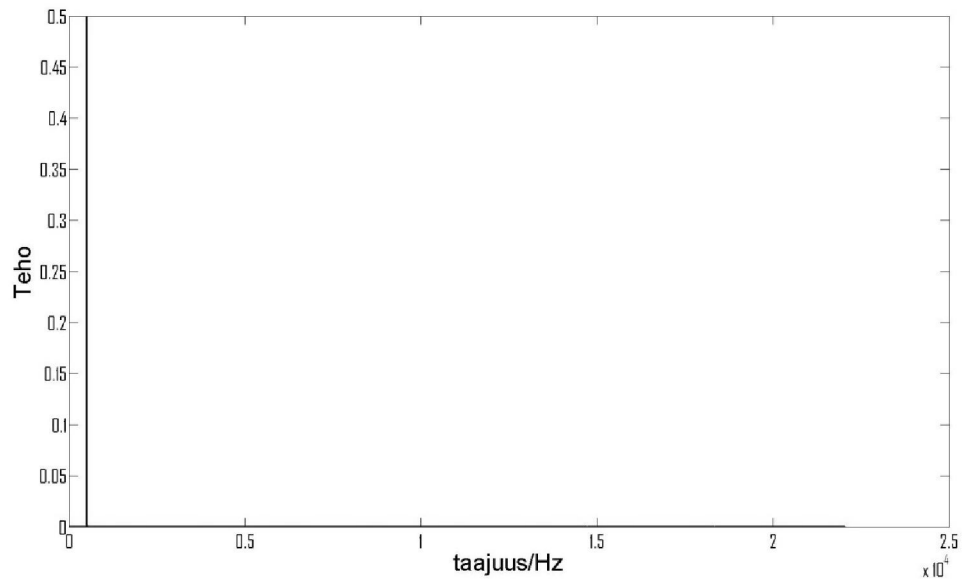
$$y = \sin(2\pi ft) \quad (52)$$

Kuvassa 7 on esitetty sinimuotoinen signaali, jonka taajuus on 500Hz.



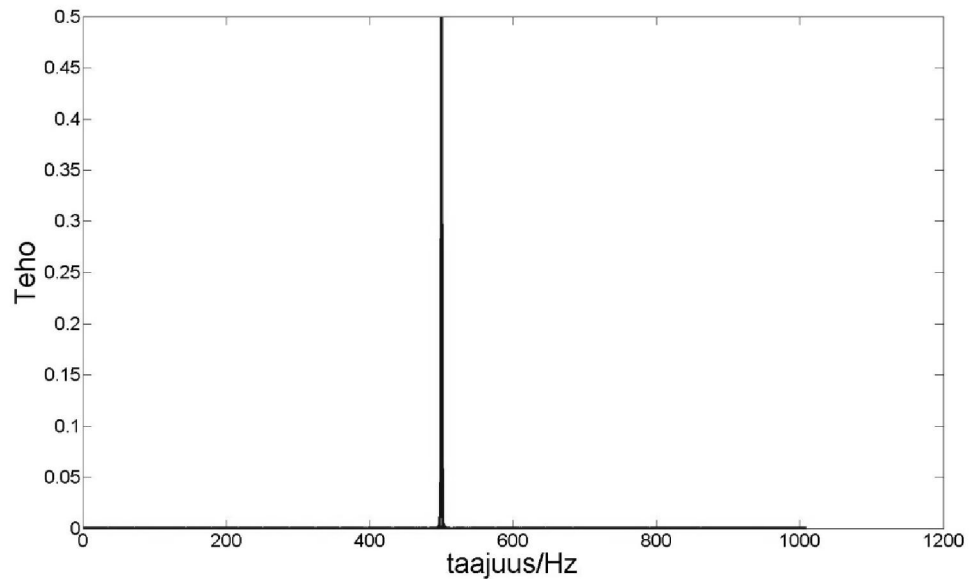
Kuva 7 Sinimuotoinen signaali, taajuus 500 Hz, näytettynä aikavälillä 0 ja 0.01

Kaavasta (5) ja (6) näemme, että signaalit voidaan jakaa taajuuskomponentteihin Fourierin muunnoksella ja näistä kaavoista näemme myös signaalin amplitudin kyseisellä taajuudella. Amplitudilla tarkoitetaan spektrin kohdalla signaalin energiasta kyseisellä taajuudella. Kuvan 7 signaali on luotu Matlab- algoritmilla. Signaalin taajuusspektri on kuvassa 8.



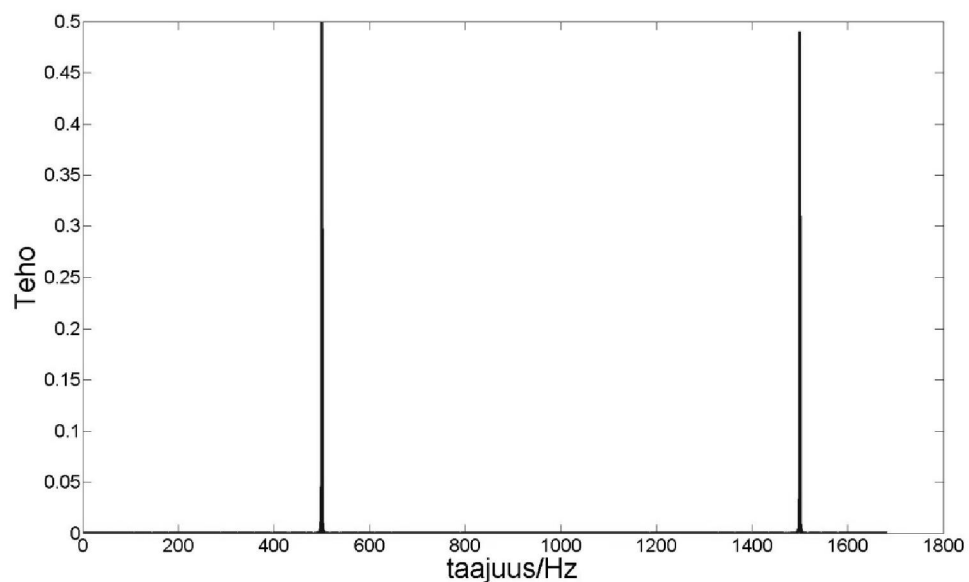
Kuva 8. Sinimuotoisen signaalin taajuusspektri FFT:llä laskettuna. Eniten energiaa on taajuudella 500Hz.

Kuvassa maksimitaajuus on 22050Hz ja signaali on taajuudella 500Hz. Lähikuva 500Hz alueesta kuvassa 9 selventää tilannetta paremmin.



Kuva 9. Sinisignaalin taajuusspektri 500Hz alueelta.

Signaalin taajuusspektristä nähdään, että signaalin taajuus näkyy hyvin selkeästi taajuusspektrissä. Mikäli voidaan olettaa, että signaali on sinisignaalin kaltainen, voidaan signaali tunnistaa etsimällä suurienergisin taajuus. Ehtona on se, että signaalissa ei esiinny muita taajuushuippuja. Kuvassa 10 on esitetty signaali missä on kaksi sinimuotoista signaalia, joista toinen taajuudella 500Hz ja toinen 1500Hz.



Kuva 10. Sinisignaalin taajuusspektrin lähikuva. Signaalissa on kaksi komponenttia 500Hz ja 1500Hz.

Kuvasta nähdään, että vaikka taajuuksilla ei aivan samaa tehoa olekaan, on ero niin pieni, että käytännössä signaalin tunnistaminen pelkästään tämän perusteella voi olla hankalaa. Tässä tilanteessa pitää muistaa myös, että molemmat esimerkki signaalit ovat hyvin yksinkertaisia ja selkeitä. Signaaleissa on hyvin vähän taajuuskomponentteja.

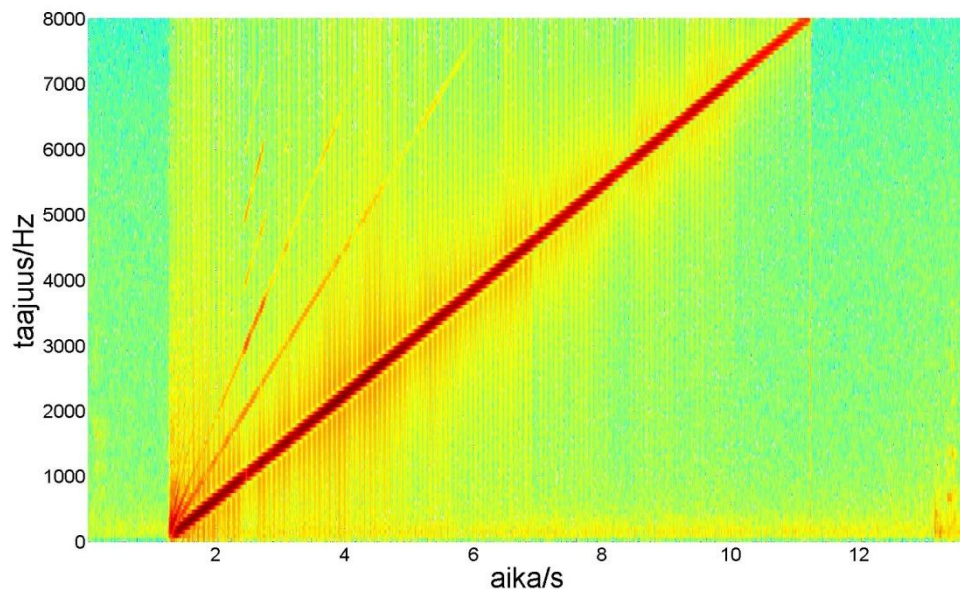
Äänisignaaleissa on yleensä useita komponentteja ja signaalien spektrit eivät aina ole näin siistejä.

3.6 Signaalin tunnistaminen spektrogrammista

Signaalinkäsittelyssä ei aina riitä, että saadaan signaalin spektri koko signaalin mitalta. Tästä näkee kyllä signaalin kaikki taajuudet mutta ei sitä, miten signaalin taajuudet muuttuvat ajan suhteen. Aina signaali ei ole samanlainen joka hetki ja lyhyen ajan taajuusspektristä voidaan saada tietoa, mitä ei koko signaalin taajuusspektristä saada ulos. Yksi tapa visualisoida taajuuden muutoksia ajan suhteen on spektrogrammi.

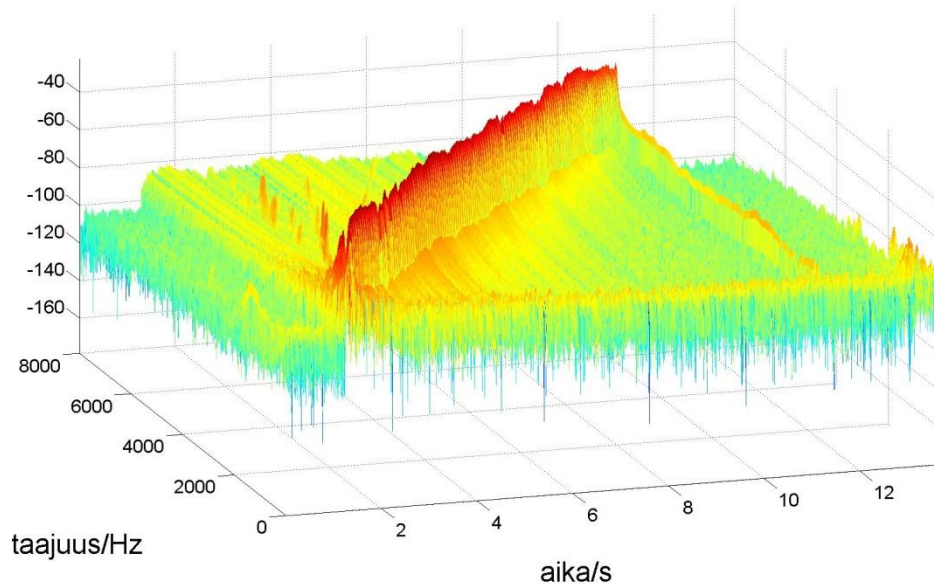
3.6.1 Spektrogrammi

Spektrogrammi kuvaa siis signaalin taajuuksien muutosta ajan suhteen. Kuvassa 11 on esimerkki sinisignaalin spektrogrammista.



Kuva 11. Sinimuotoisen signaalin spektrogrammi. Signaalin taajuus nousee tasaisesti koko 1 hertsistä 8000 hertsiin.

Signaalin taajuus nousee 10 sekunnissa 1Hz 8000Hz. Signaali on luotu kaavan (52) mukaisesti. Vaaka-akselilla on aika sekunneissa ja pystyakselilla on taajuus hertseissä. Kunkin taajuuden teho kuvataan väreillä. Punaisella alueella on eniten tehoa ja vihreällä vähiten. Spektrogrammi voidaan esittää myös kolmiulotteisena kuten kuvassa 12.



Kuva 12. Spektrogrammin kolmiulotteinen esitystapa.

Kuvan 12 punainen harju vastaa kuvan 11 punaista viivaa. Molemmat kuvat ovat saatu Matlab-ohjelmiston ”spectrogram” rutiinilla

Spektrogrammin laskemista varten tarvitaan niin sanottu lyhytaikainen Fourierin muunnos eli STFT (Short-Time Fourier Transform). Tämä on määritelty seuraavalla tavalla

$$X_{STFT}(e^{j\omega}, n) = \sum_{m=-\infty}^{\infty} x[n-m]w[m]e^{-j\omega m} \quad (53)$$

(Mitra 2006), missä $w[m]$ on ikkuna, jonka tarkoituksena on äärellisen mittainen osa signaalista tarkasteluun. Tuon ikkunan aikana signaalin spektrin olisi hyvä olla mahdollisimman stationaarinen. (Mitra 2006). STFT:lla saatu tulos on kahden muuttujan funktio, aika ja taajuus. Yleensä kiinnostavin tieto STFT:sta on sen magnitudi. Tämän magnitudin visualisointia kutsutaan yleensä spektrogrammiksi. (Mitra 2006). Eli kyseessä on signaalin spektri peräkkäisillä lyhyillä ajanjaksoilla. Nämä signaalin muodostavat ikään kuin kolmiulotteisen matriisin, mihin yhdelle riville on talletettu signaalin spektri tietyllä ajankohdalla. Koska STFT:ssa on jo edellä mainitut kaksi muuttujaa (aika ja taajuus), sen esittäminen ei ole aivan yksinkertaista. Yleensä sitä visualisoidaan kuvien 11 ja 12 tapaan. Näissä kuvissa magnitudi on värikoodattu siten, tumman punainen on suurin magnitudi ja vihreästä siniseen on pienin. Värikoodauksen voi tehdä miten haluaa ja se riippuu täysin sovelluksesta. Kuvassa 12 on kolmiulotteinen visualisointi, mihin on liitetty mukaan värikoodaus. Spektrogrammia käytetään teollisuudessa muun muassa musiikkikappaleiden tunnistamiseen. Tästä on oivana esimerkkinä Shazam (Wang 2003).

3.7 Äkkinäiset muutokset signaalissa

Tässä työssä yksi ensimmäisistä haasteista oli keksiä keino, millä voidaan havaita signaalissa esiintyvä kova ääni. Kovalla äänellä tarkoitetaan signaalin amplitudissa ja tehossa tapahtuvaa äkkinäistä muutosta, joka voi olla pysyvä tai lyhytaikainen häiriö. Näiden havaitseminen voi olla erittäin tärkeää, sillä pienet häiriöt ja poikkeamat signaalissa saattavat häiritä mittauksia ja tuottaa ei-toivottuja tuloksia. Äkkinäinen muutos tulkitaan tässä työssä audiosignaalien kohdalla signaalin amplitudin voimakkaaksi muutokseksi eli toisin sanoen signaalin desibelitason äkkinäiseksi nousuksi. Äkkinäistä muutosta voidaan havaita signaalissa kahdella tavalla. On-line menetelmällä yritetään havaita signaalissa tapahtuva muutos tuntemattomana ajankohta eli periaatteessa yritetään havaita muutos reaaliajassa tai hyvin lyhyessä ajassa. Off-line menetelmissä koko signaali on käytössä ja sitä analysoimalla pyritään löytämään äkkinäisen muutoksen kohta signaalissa. (Jokinen 2003).

3.8 Signaalin energia ja teho

Signaali energia ja teho ovat määritelmällisesti eri asioita. Signaalin $x[n]$ energia on

$$E_x = \sum_{m=-\infty}^{\infty} |x[n]|^2 \quad (54)$$

(Mitra 2006). Tästä seuraa, että joillakin signaaleilla voi olla ääretön energia. Mikäli signaali $x[n] = \frac{1}{\sqrt{n}}, n \geq 1$, tästä seuraa kaavan (54) mukaan, että $E_x = \sum_{n=1}^{\infty} \frac{1}{n}$. Tämä ei suppene mihinkään arvoon, vaan saadaan signaalin energiaksi ääretön. On tietenkin olemassa signaaleita, joilla on äärellinen energia. Energian avulla voidaan laskea signaalille keskimääräinen teho. Tehon laskeminen aloitetaan määrittelemällä energia ensiksi tietyn intervallin yli. (Mitra 2006) Tätä varten muokataan kaavaa (54) hiukan

$$E_{x,K} = \sum_{m=-K}^K |x[n]|^2 \quad (55)$$

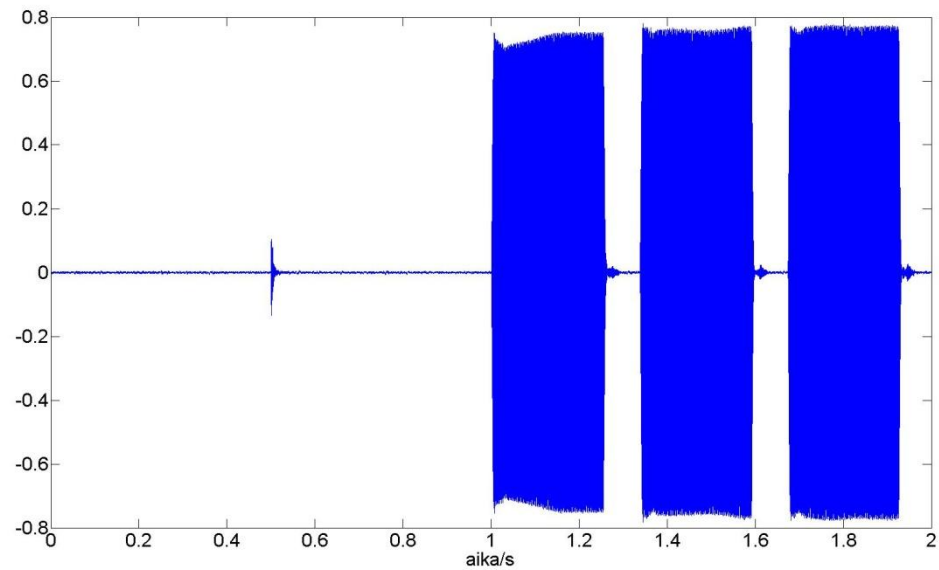
(Mitra 2006). Intervalli on tässä tapauksessa $-K \leq n \leq K$. Koska teho on energia per aikayksikkö (Young 2008) tai tässä tapauksessa energia per näyte, voidaan signaalin teho esittää signaalin keskimääräisenä energiana

$$P_x = \frac{1}{N} \sum_{m=-K}^K |x[n]|^2 \quad (56)$$

(Mitra 2006). Kaavassa (56) $N=2K+1$ eli ikkunan pituus. Signaalin teho voi olla joko äärellinen tai ääretön, kuten signaalin energiakin. Signaalit voidaan jakaa energian ja tehon määritelmien avulla kahteen eri kategoriaan; tehosignaalit ja energiasignaalit. Energiasignaaleilla on äärellinen energia mutta ääretön teho kun tehosignaaleilla on äärellinen teho ja ääretön energia. (Mitra 2006). Joissakin sovelluksissa on hyödyllistä jakaa signaali osiin ja laskea sieltä energia kaavalla (54) ilman, että lasketaan tehoa. Signaalissa tapahtuvaa äkkinäistä muutosta voi laskea siis tehosta ja energiasta.

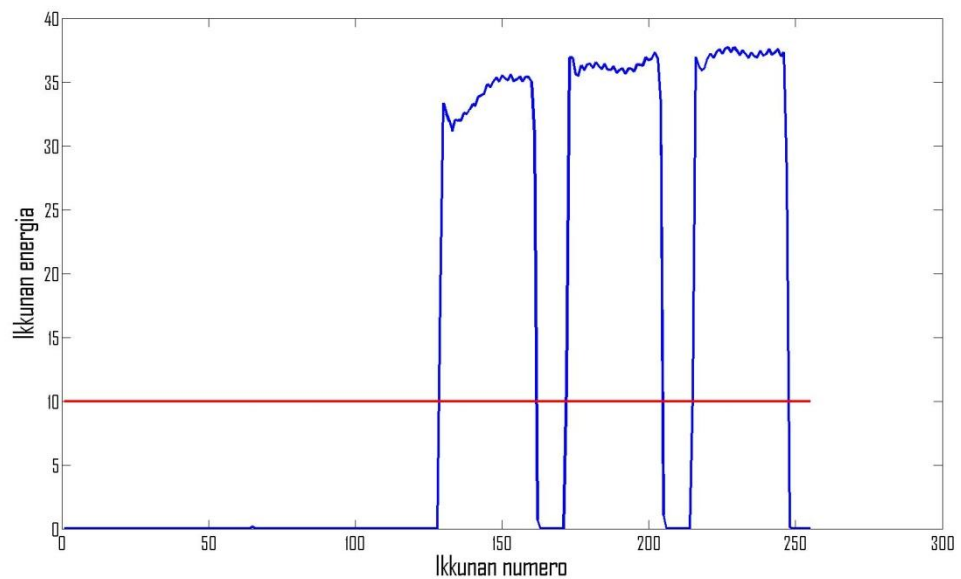
3.8.1 Muutoksen havaitseminen energiasta ja tehosta

Äkkinäisen muutoksen havaitseminen signaalissa sen tehon tai energian kautta on suhteellisen yksinkertaista. Ei ole mielekästä laskea signaalin energiaa jokaiselle näytteelle vaan pikemminkin jakaa signaali tasamittaisiin ikkunoihin, joista lasketaan energia ja näitä ikkunoita vertaillaan toisiinsa. Tällaisen vertailun voi tehdä joko jälkikäteen tai sitten reaaliajassa. Signaalin energialle on annettava joku kynnysarvo, minkä jälkeen oletetaan, että signaalissa on tapahtunut äkkinäinen muutos. Kynnysarvo riippuu hyvin paljon sovelluksesta ja halutusta tuloksesta. Yksinkertaisimmillaan voidaan todeta, että kun signaalin energia kasvaa x -kertaiseksi edelliseen ikkunaan nähden, oletetaan, että äkkinäinen muutos on tapahtunut. Kuvassa 13 on esitetty normaalin asuinrakennuksiin tarkoitetun palovaroittimen kolme ensimmäistä piipahdusta aikatasossa. Kuvassa ensimmäinen äännähdyn nähdään amplitudin kasvamisena yhden sekunnin kohdalla. Sitä ennen on nauhoitettu huoneesta taustamelua, joka ei ole juuri muuta kuin kohinaa. Kuvassa 11 on tämän signaalin energia. Energia on laskettu ikkunoittain ja kuvaajassa näkyy kunkin ikkunan, joita on 255 kappaletta, energia. Kuvasta 14 nähdään myös, että ikkunan 129 energia nousee kynnysarvon yli ja tämä tulkitaan äkkinäiseksi muutokseksi signaalissa eli kovaksi ääneksi, koska kyseessä on audiosignaali. Kuvassa 14 oleva kynnys on siis absoluuttinen ja siinä vertaillaan vain kunkin ikkunan omaa arvoa, riippumatta muiden ikkunoiden energioista. Mikäli ikkunan energiaa verrataan aikaisempiin ikkunoihin, ei kynnystä tarvitse määrittää. Kertoimen x ja kynnysarvon määrittely on tässä työssä heuristinen prosessi ja sen saama arvo voi vaihdella hyvin paljon.



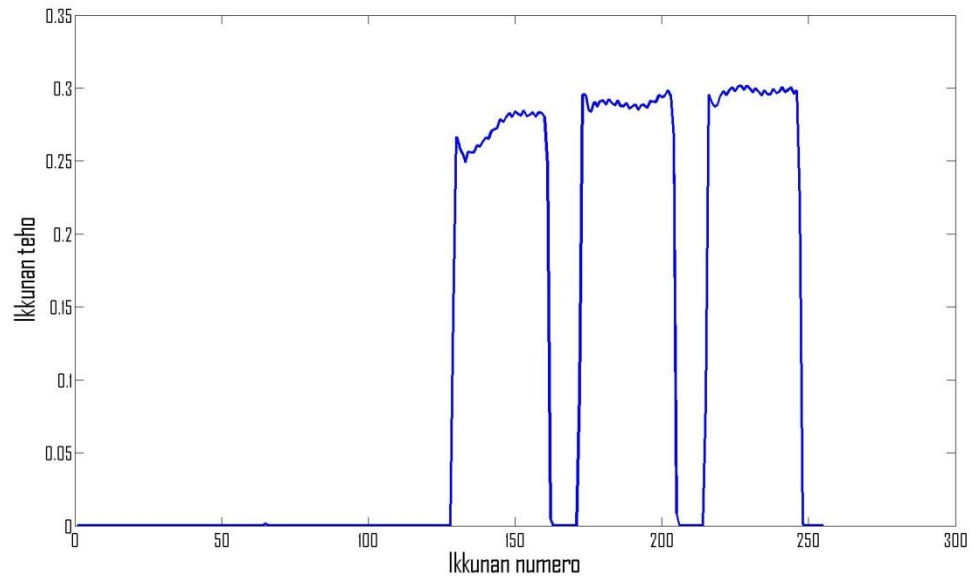
Kuva 13. Palohälyttimen kolme ensimmäistä piippausta kuvattuna aikatasossa. Näytteenottotaajuutena 16000 Hz.

Kuva 13 on tehty käyttämällä Matlabin plot-komentoa.



Kuva 14. Kuvassa on kuvan 13 signaalin energia ikkunoittain. Kunkin ikkunan energia on laskettu kaavalla (55). Punaisella on merkitty yksi mahdollinen kynnyсарvo, mihin signaali voidaan verrata.

Kuvassa 14 on käytetty signaalista otetun ikkunan energiaa, mutta voidaan käyttää myös kyseisen ikkunan tehoa. Teho lasketaan kaavalla (56) ja se on siis energia per näyte. Kuvassa 15 on esitetty signaalin teho ikkunoittain.



Kuva 15. Kuvassa on kuvan 13 signaalin teho ikkunoittain.

Huomataan, että kynnysarvo on laskettava uudelleen, koska tehon absoluuttiset arvot ovat alle yhden. Tässäkin tapauksessa kuitenkin ikkunan 129 keskimääräinen teho on huomattavasti suurempi kuin edeltävän ikkunan. (Dufaux 2001).

Kynnysarvon määrittämisessä voidaan ottaa huomioon myös useampien edellisten ikkunoiden tehon keskiarvo. Signaalissa voi esiintyä yksittäisiä suhteellisen kovia ääniä, jotka eivät ole hälytyksiä mutta nousevat silti absoluuttisen kynnysarvon ylitse. Näiden vaikutus pienenee, kun käytetään usean edellisten ikkunoiden tehojen keskiarvoa kynnysarvona. Tällöin kynnysarvo on dynaaminen ja siihen vaikuttaa ympäröivä melu. Järjestelmä osaa tällöin adaptoitua ympäristöönsä.

4 HAHMONTUNNISTUS

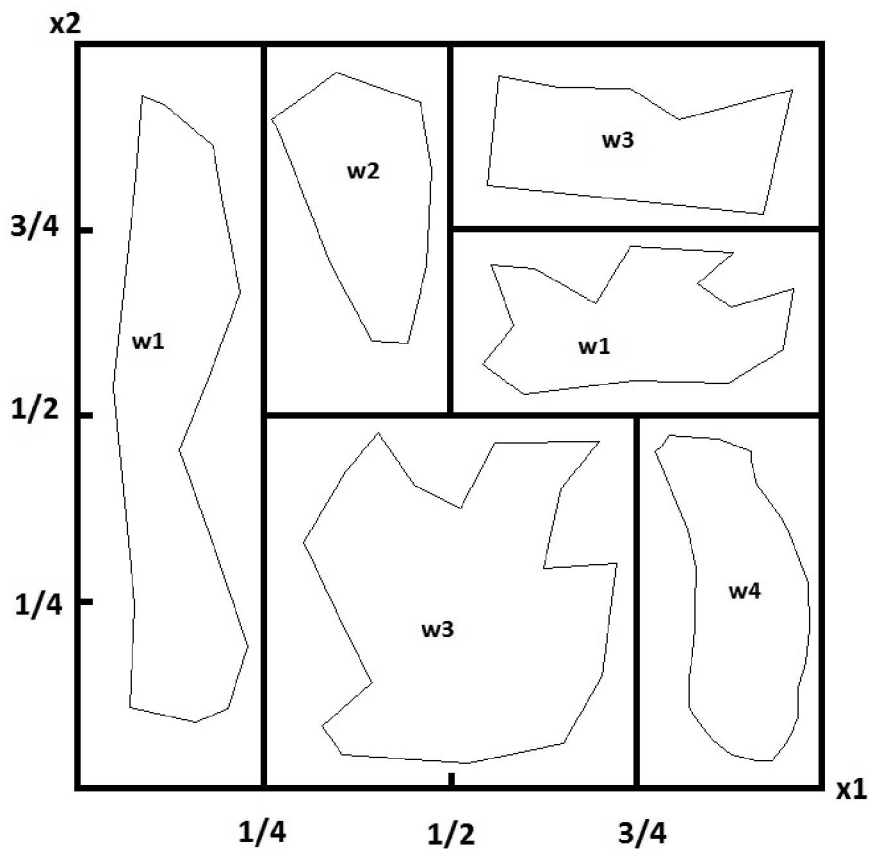
Hahmontunnistus on periaatteessa jonkin objektin sijoittamista johonkin luokkaan tiettyjen kyseisestä objektista mitattujen piirteiden perusteella (Koutroumbas 2006). Jokaisessa luokassa on samanlaiset piirteet omaavat objektit. Mitattavien piirteiden määrä voi vaihdella miten paljon vain ja tämä tekee jokaisesta hahmontunnistusoperaatiosta omanlaisensa. Piirteitä on oltava riittävän paljon, että luokitus voidaan tehdä luotettavasti mutta silti riittävän pieni, että piirteiden hakuun ja niiden vertailuun ei mene aikaa tai laskentatehoa liikaa. Tästä seuraa pieni tasapainoilu tarkkuuden ja ajan välillä. Optimaalisin tilanne olisi sellainen, missä muutamalla piirteellä voitaisiin luokitella suhteellisen varmasti jokainen objekti.

Hahmontunnistus voidaan jakaa viiteen eri vaiheeseen: 1. Mittaus; 2. Esikäsittely eli segmentointi; 3. Piirteenirrotus; 4. Luokitus; 5. Jälkikäsittely. Mittauksessa objektista hankitaan dataa, esimerkiksi ottamalla valokuva tai nauhoittamalla ääntä. Esimerkiksi voidaan nauhoittaa henkilön puhetta ja yrittää tunnistaa mitä hän sanoi tai ottaa valokuva eläimen jalanjäljistä. Esikäsittelyssä saatua dataa voidaan suodattaa tai sille voidaan tehdä muita operaatiota. Tarkoituksena on tuoda haluttuja piirteitä paremmin esille mahdollisesti hyvin suuren kohinan seasta. Segmentoinnissa data pyritään jakamaan yhden luokiteltavan objektin kokosiin osiin. Äänisignaaliissa signaali jaetaan yhden kirjaimen mittaisiin osiin tai valokuvasta erotetaan yksi jalanjälki. Piirteenirrotuksessa datasta otetaan halutut piirteet esiin. Yleensä informaatiota on liikaa ja tässä vaiheessa sitä pyritään supistamaan niihin osiin, joilla tunnistaminen onnistuu varminten. Näin saadaan turhat piirteet pois ja saatavaa dataa on helpompi käsitellä. Äänisignaalin kohdalla tämä voi olla signaalin taajuus tai sen voimakkuus, kuvasta voidaan yrittää mitata jalanjäljen pituus ja leveys. Näistä irrotettavista piirteistä saadaan piirvektori, joka annetaan luokittimelle. Luokitin suorittaa luokituksen eli sijoittaa objektin annettujen piirteiden perusteella johonkin luokkaan. Pyrkimys on, että mahdollisimman moni objektista luokitetaan oikein. Jälkikäsittelyllä tarkoitetaan lopullista toimenpidettä, mikä pitää tehdä luokituksen jälkeen. Mikäli valokuvasta tiedetään olevan vain yhden eläimen jälkiä ja niistä yksi luokitellaan suden jäljiksi ja loput karhun, voidaan olettaa, että kyseessä on karhu. (Tohka 2012).

Hahmontunnistus voidaan tehdä kahdella tavalla: ohjatusti tai ohjaamatta. Ohjatussa oppimisessa tiedetään luokat etukäteen ja tiedetään millaisia piirvektoreita niihin kuuluu. Tällöin on käytössä ollut jokin oppimisdata, minkä perusteella luokat on valittu. Hahmontunnistusjärjestelmän on tarkoitus etsiä näistä luokista sopivin. Ohjaamattomas-

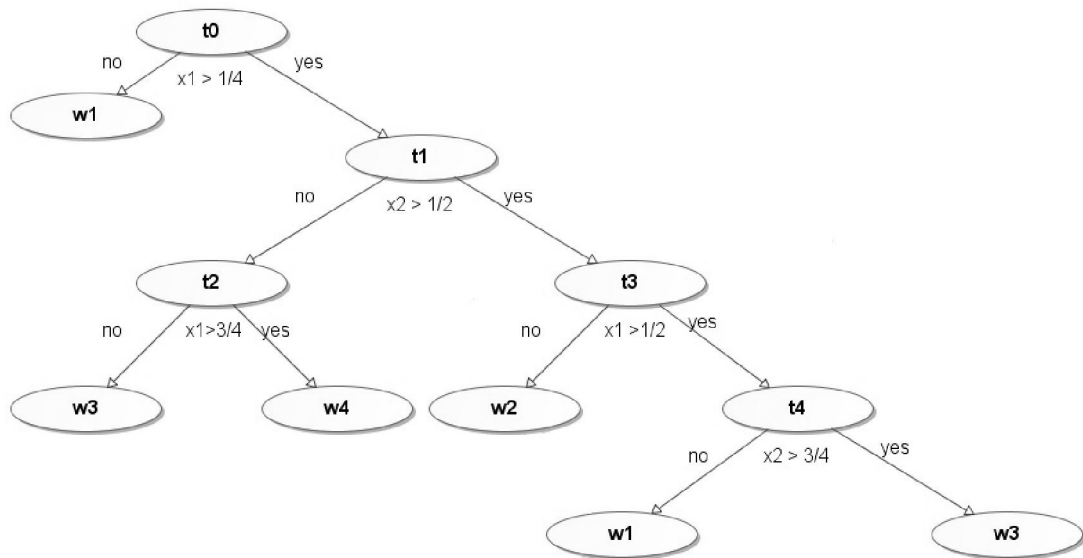
sa oppimisessa ei oppimisdataa ole käytössä vaan järjestelmän on etsittävä datasta samankaltaiset piirteet omaavat objektit ja luokiteltava ne itse. Tällöin puhutaan klusteriionnista. (Koutroumbas 2006). Kolmaskin tapa on olemassa, niin sanottu vahvistusoppiminen. Tässä ei ole käytössä myöskään oppimisdataa mutta järjestelmälle kerrotaan onko sen tekemä luokitus oikein vai väärin. Ohjattu ja ohjaamaton oppiminen ovat kuitenkin ehdottomasti yleisimmät tavat. (Tohka 2012). Tämän työn puitteissa puhutaan ennen kaikkea ohjatusta oppimisesta. Ennen kuin hahmontunnistuksen viisi vaihetta suoritetaan, on etukäteen jo opetettu mitkä luokat tunnetaan ja mitkä niiden piirteet ovat.

Hahmontunnistuksessa on tärkeää se miten, eri piirteitä käytetään tunnistamiseen. Tähän työhön valittiin valintapuu- menetelmä (decision tree) (Koutroumbas 2006). Valintapuu on ei-lineaarinen menetelmä, missä eri luokat hylätään yksi kerrallaan kunnes jäljellä on vain oikea luokka. Valintapuussa jokaista piirrettä verrataan tiettyyn kynnyksarvoon ja sen mukaan suoritetaan luokitusta. Kuvassa 16 on esitetty kuva miten eri piirteet ovat jaettu piirreavaruudessa. Kuvasta huomataan, että koska koko piirreavaruus on jaettu luokkiin, ei tuntemattomia luokkia voi olla.



Kuva 16. Valintapuuta varten on tehty piirreavaruuden jaottelu.

(Koutroumbas 2006). Kuvassa 17 on valintapuu. Siitä käy hyvin selville miten eri piirteet sijoitetaan eri luokkiin.

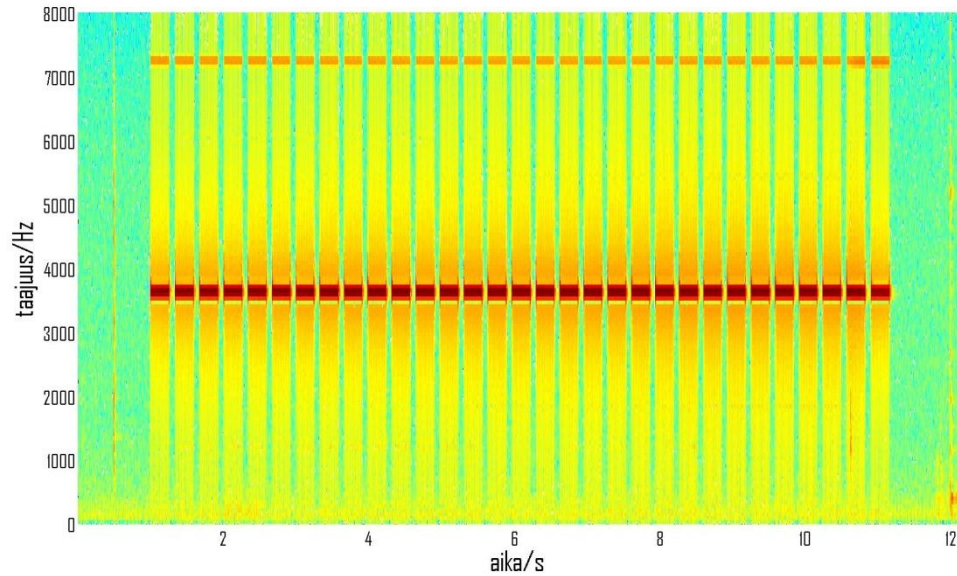


Kuva 17. Päätöspuussa tapahtuvan luokituksen periaatteet.

(Koutroumbas 2006). Kuvan 17 oleva puu on binaarinen. Siinä merkintä "tx" merkitsee kohtaa, missä tarkkaillaan piirteen x_1 tai x_2 arvoa verrattuna johonkin kynnysarvoon. Tätä kutsutaan myös solmuksi. Kohdassa t_0 tarkastellaan onko $x_1 > 1/4$. Kynnysarvo löydetään kuvasta 16 vaaka-akselilta. Mikäli x_1 on pienempi kuin $1/4$, luokitellaan se luokkaan w_1 kuvan 16 mukaisesti. Mikäli x_1 on suurempi kuin $1/4$, jatketaan seuraavaan solmuun t_1 . Tässä tarkastellaan onko $x_2 > 1/2$ eli tarkastellaan toista piirrettä. Mikäli x_2 on pienempi kuin $1/2$ tarkastellaan onko arvo $x_1 > 3/4$. Kuvasta 16 nähdään, että mikäli x_2 on pienempi kuin $1/2$, niin luokittelu tapahtuu piirteen x_1 mukaan. Mikäli $x_1 > 3/4$, niin luokaksi muodostuu w_4 , muussa tapauksissa luokka on w_3 . Tässä pitää muistaa, että tilanne missä $x_1 < 1/4$, luokiteltiin jo solmussa t_0 . Näin jatketaan loppuun asti. Kuvasta 17 huomataan, että joka tilanteessa saadaan jokin luokka määrättyä. (Koutroumbas 2006).

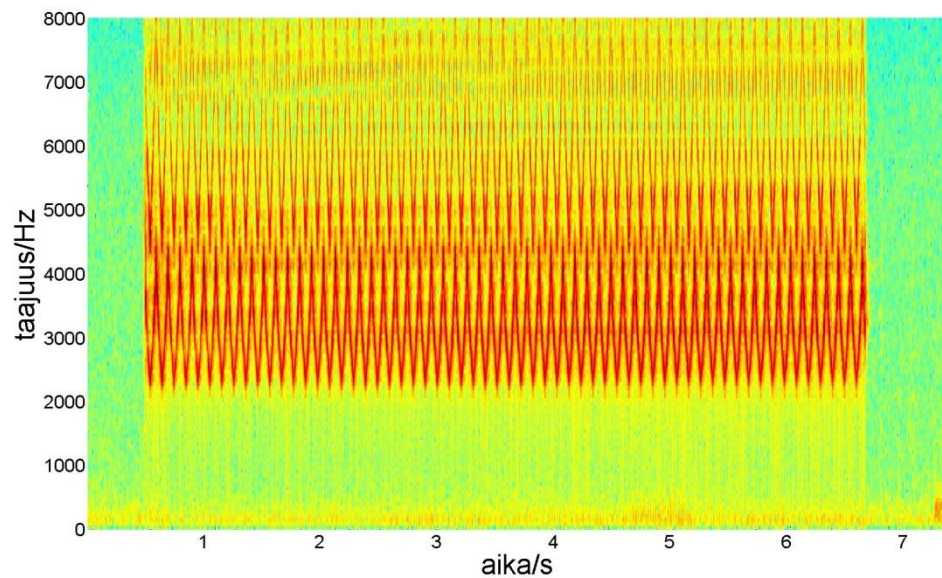
4.1 Hälyttimien äänen analysointi

Tässä diplomityössä on keskitytty kahteen ääneltään erilaiseen hälytin tyyppiin. Ryhmässä 1 ovat tavallisessa kotikäytössä olevat palohälyttimet, savuvaroittimet, kosteushälyttimet, kaasuhälyttimet ja niin edelleen, joista tulee piipittävä ääni. Ryhmässä 2 erikoiskohtaiset hälyttimet, murtohälyttimet ja niin edelleen, joiden taajuus vaihtelee koko ajan. Ryhmässä 2 ei ole lainkaan taukoja hälytyksessä kuten ensimmäisessä on. Kuvassa 18 on palohälyttimen spektrogrammi.



Kuva 18. Palohälyttimen spektrogrammi, nauhoitettu kaiuttomassa huoneessa.

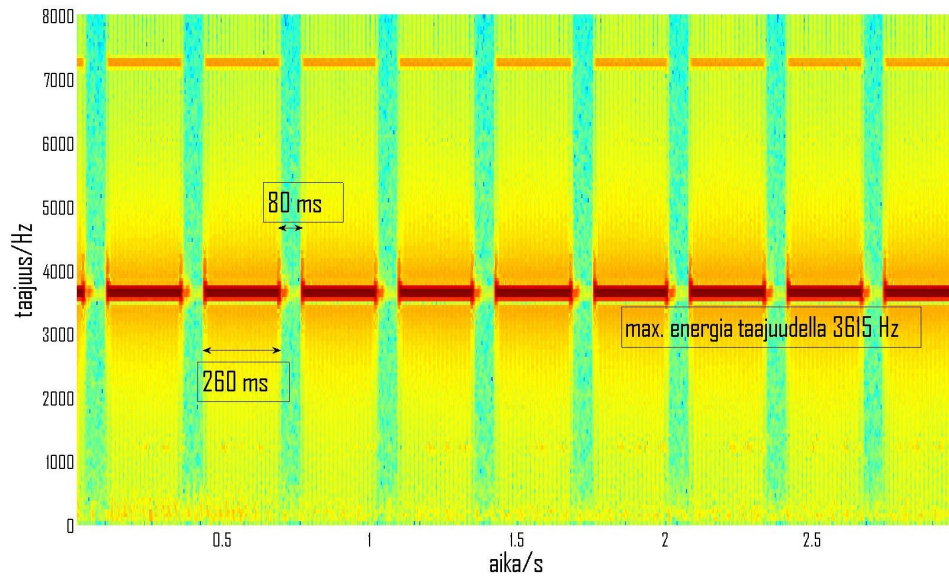
Siitä nähdään hälyttimen äänen luonne. Lyhyitä ääniä noin 3500 Hz alueella. Muualla ei merkittäviä taajuuskomponentteja ole. Kaikkien hälytinten taajuus ei ole sama mutta kaikkien ovat noin 3000 Hz ja 4000 Hz välissä. Palovaroittimille on olemassa omat standardit, jotka määrittelevät niiden taajuuden, kuten kappaleessa 2 kerrottiin. Kuvassa 19 on erikoishälyttimen spektrogrammi.



Kuva 19. Erikoishälyttimen spektrogrammi. Kuvasta havaitaan ryhmän 2 hälyttimille tyypillinen sahalaita kuvio ja sen kerrannaiset.

Tässä huomataan ero ryhmän 1 hälyttimiin. Äänen taajuus vaihtelee sahalaitakuvion mukaisesti eikä äänessä ole taukoja. Ryhmän 2 hälytinten äänet vaihtelevat noin 2000 Hz ja 5000 Hz välillä. Äänet ovat selkeästi erilaisia. Molemmissa kuvissa näkyy päätaajuuksien kerrannaisia eli harmonisia taajuuksia.

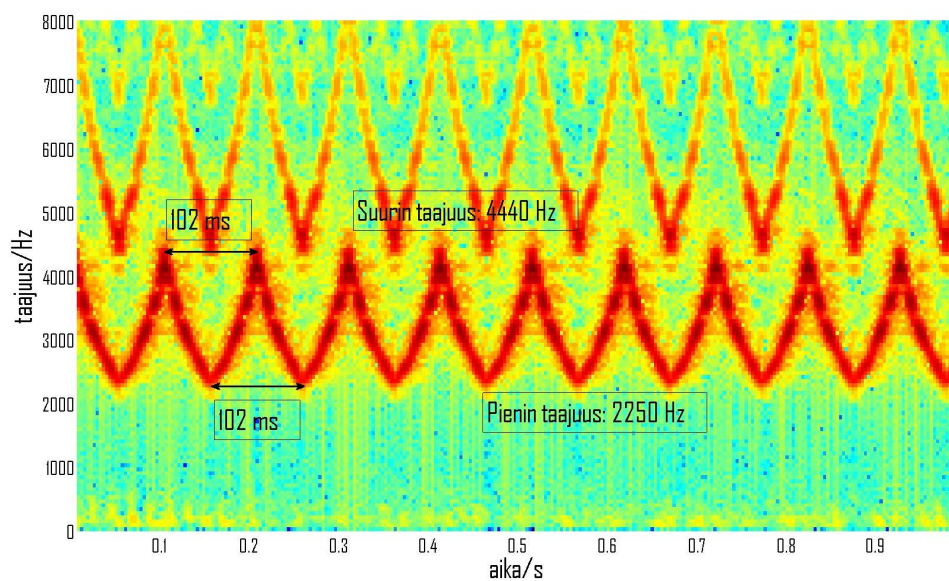
Ryhmän 1 hälyttimissä on myös muita eroja. Äänen kesto ja hiljaisuuden kesto ajassa vaihtelevat hälytinten välillä. Kuvassa 20 on edellä mainitun palohälyttimen spektrogrammin lähikuva.



Kuva 20. Palohälyttimen spektrogrammi. Kuvaan on merkattu äänen ja hiljaisuuden kesto sekä maksimienergian taajuus.

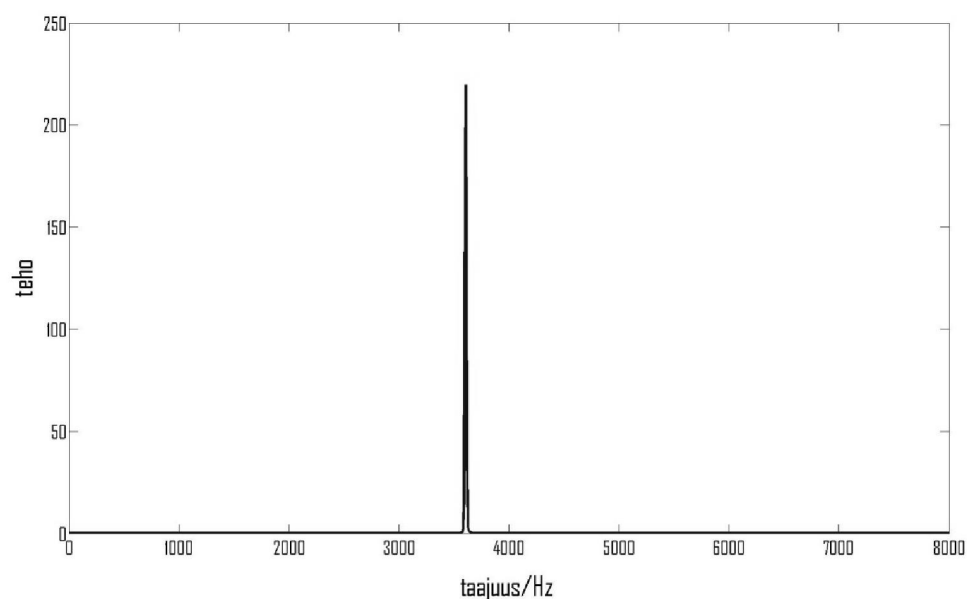
Kuvasta näemme, että signaalin äänen eli yhden piipin kesto on noin 260 millisekuntia ja piipin välisen hiljaisuuden on 80 ms. Suurin taajuuskomponentti on noin 3600 Hz kohdalla.

Samankaltaiset eroavaisuudet voidaan löytää myös ryhmän 2 hälyttimien kesken. Kuvassa 21 on erikoishälyttimen spektrogrammi.

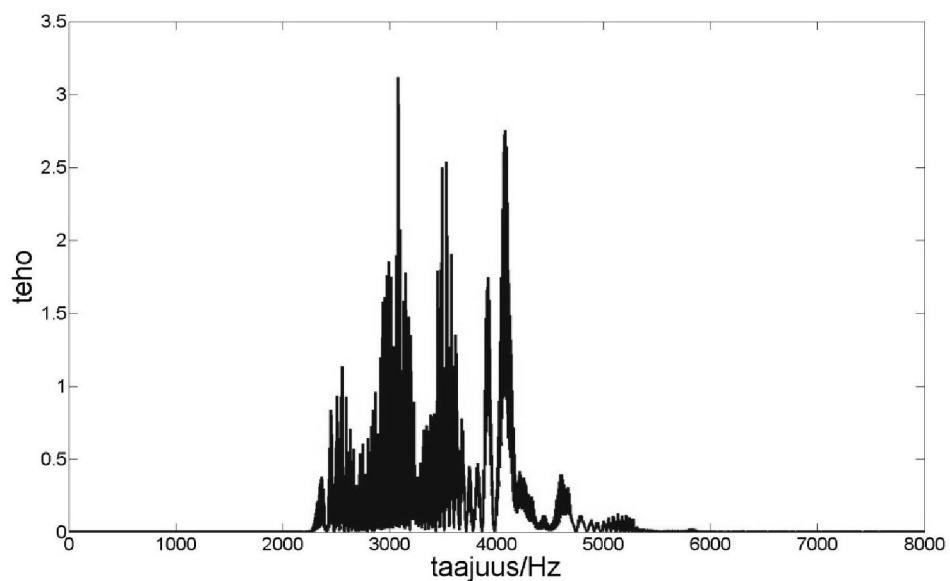


Kuva 21. Erikoishälyttimen spektrogrammi.

Äänen sahalaidan maksimi- ja minimitaajuudet ovat näkyvissä kuvassa. Ne ovat 4800 Hz ja 2250 Hz. Näiden lisäksi voidaan mitata sahalaidan huippujen ja kestot eli tois sanoen jaksonaika. Kaikilla ryhmän 2 hälyttimillä tämä on säännöllinen ja erikoishälyttimellä se on 102 ms. Eri hälytinryhmien välillä on selkeä ja havaittavissa oleva ero ja ryhmän sisällä signaaleilla on eroavaisuuksia, joita voidaan käyttää hyväksi signaaleita tunnistessa. On hyvä huomata, että ryhmän 1 hälytinten teho on jakautunut hyvin pienelle taajuuskaistalle ja ryhmän 2 energia on laajemmalla alueella. Tämän näkee hyvin, kun otetaan signaalin tehospektri pidemmältä ajanjaksolta. Kuvassa 22 on ryhmään 1 kuuluva palohälyttimen 1 spektri ja kuvassa 23 on ryhmään 2 kuuluva erikoishälyttimen spektri.



Kuva 22. Palohälytin 1:n spektri. Hälytin kuuluu ryhmään 1.



Kuva 23. Erikoishälyttimen spektri. Hälytin kuuluu ryhmään 2.

Molempien spektri on yhden sekunnin mittaisesta näytteestä joten molemmissa on kaikki ne taajuudet mitä hälyttimen ääneen kuuluu. Huomataan, että ryhmän 1 hälytin kuvassa 22 on teholtaan huomattavasti kapeammalla alalla kuin ryhmään 2 kuuluva hälytin kuvassa 23. Tämä on yksi huomattava ero näiden kahden ryhmän välillä. Kuvat 22 ja 23 ovat hyvin tyypillisiä ryhmiensä sisällä ja kaikilla hälyttimillä on suurin piirtein samanmuotoiset spektrit. On hyvä huomata myös, että maksimiteho on huomattavasti suurempi ryhmän 1 hälyttimillä kuin ryhmän 2.

4.2 Hälyttimet ja niiden ominaisuudet

Ryhmän 1 hälyttimistä on esimerkkinä tässä työssä palohälytin 1 ja ryhmästä 2 erikoishälytin. Alun perin äänien ominaisuudet laskettiin silmämääräisesti spektrogrammin avulla ja vasta myöhemmin tehtiin Matlab- algoritmi, millä kyseiset arvot laskettiin. Laskennalliset arvot ja mitatut arvot poikkeavat hiukan toisistaan mutta mitään suuria muutoksia näiden välillä ei ole. Taulukoissa 1 ja 2 on koottu eri hälyttimien ominaisuudet.

Palohälytin 1: Kuvissa 18, 20 ja 23 on palohälytin 1 äänen spektrogrammit ja taajuuspektri, joista äänen ominaisuudet ovat havaittavissa. Palohälytin 1 paristoilla toimiva normaaliin kotikäyttöön tarkoitettu palovaroitin, jonka toiminta perustuu savun havaitsemiseen. Tämän hälyttimen toiminta ja ominaisuudet ovat määritelty standardissa SFS-EN-14604, joka käsiteltiin kappaleessa 2. Hälyttimen ääni ei ole jatkuva vaan siinä toistuu äänellinen jakso ja äänetön jakso. Äänellinen jakso kestää noin 260 ms ja äänetön jakso noin 80 ms eli yhteensä yhden jakson kesto on 340 ms. Nämä ajat pysyvät vakiona koko hälytyksen ajan. Hälyttimen taajuus on hiukan yli 3600 Hz ja sekin pysyy vakiona hälytyksen aikana. Kuvista voidaan laskea äänen toistumistiheys. Palohälyttimellä 1 tämä on noin 3 Hz. Tämä voidaan varmistaa laskemalla taajuus äänen jaksonajasta, joka on siis äänellisen jakson ja äänettömän jakson kesto eli $260\text{ms} + 80\text{ms} = 340\text{ms}$. Taajuus $f = \frac{1}{T} = \frac{1}{340\text{ms}} = 2.94$ eli noin 3 Hz. Tämäkin arvo pysyy vakiona koko hälytyksen ajan. Matlab- algoritmilla saadut arvot ovat äänellinen jakso: 245 ms, äänetön jakso: 85 ms, taajuus 3615 Hz ja äänen toistumistiheys 3 Hz. Tulokset olivat siis hyvin lähellä käsin mitattuja tuloksia.

Erikoishälytin: Kuvista 19, 21 ja 23 nähdään erikoishälyttimen ominaisuudet. Kuvasta 21 nähdään erikoishälyttimelle ja koko ryhmälle 2 tyypillinen sahalaitakuvio. Tätä voidaan käyttää hälyttimen tunnistamiseen. Sama sahalaita näkyy kerrannaisina alkuperäisen sahalaidan yläpuolella. Hälyttimen ominaisuudet ovat; sahalaidan alin taajuus: 2250 Hz, sahalaidan ylin taajuus: 4440 Hz, sahalaidan jaksonaika 95 ms eli sen taajuus on $f_{\text{sahalaita}} = \frac{1}{T_{\text{sahalaita}}} = \frac{1}{102\text{ms}} = 9.8\text{Hz}$.

Taulukko 1. Hälyttimien äänten erilaisia ominaisuuksia. Matlab-sarake viittaa edellisen hälyttimen Matlabin algoritmilla saatuihin laskennallisiin tuloksiin.

	palohälytin 1	matlab
äänellinen jakso / ms	260	245
ääneton jakso / ms	80	85
maksimienergian taajuus / Hz	3615	3605
äänen toistumistiheys / Hz	2.9	3.02

Taulukko 2. Ryhmän 2 hälyttimien ominaisuuksia. Matlab-sarakkeessa on Matlab-algoritmilla saadut laskennalliset arvot.

	Erikoishälytin	matlab
sahalaidan alin taajuus / Hz	2250	2356
sahalaidan ylin taajuus / Hz	4438	4775
sahalaidan jaksonaika / ms	102	103
sahalaidan taajuus / Hz	9.8	9.7

Taulukoista 1 ja 2 nähdään, että kuvaajista arvioidut arvot vaihtelevat jonkin verran algoritmilla saaduista arvoista. Erot eivät kuitenkaan ole suuria.

5 TESTAUS

5.1 Hälyttimien testaus

Hälyttimille suoritettiin useita erilaisia testauksia eri ympäristöissä. Tällä pyrittiin selvittämään miten hälyttimet käyttäytyy eri ympäristöissä ja miten hälyttimen ääni vaimenee tai muuttuu, mikäli hälytin on vaimennettu jollain tavoin. Näiden lisäksi pyrittiin selvittämään miten eri laitteet muokkaavat ääntä, jos yhtään mitenkään.

Hälyttimen äänen vaimentumista testattiin asettamalla vastaanottimen päälle erilaisia vaatteita kaiuttomassa huoneessa, pihalla, toimistossa ja erityisen kaikuvassa tilassa. Laitteiden toimivuutta testattiin kaiuttomassa huoneessa nauhoittamalla ääntä handsfreellä, nauhoittamalla valkoista kohinaa tai nostamalla sinisignaalin taajuutta 10 sekunnin ajan.

5.2 Laitteiden aiheuttama vaimennus

Sini-sweepissä tuotetaan ääni signaali, jolla on yksi taajuuskomponentti. Tuo taajuuskomponentti vaihtelee halutulla välillä tietyssä aikana. Tässä työssä käytettiin taajuusväliä 1-8000 Hz. Tämä valittiin, koska se on nauhoittavien ohjelmien maksimi näytteenottotaajuus. Tällä tavoin voidaan testata miten vastaanottava laite reagoi tiettyyn taajuuteen tai vaimentaako jokin esine tai materiaalia jotain taajuutta enemmän kuin toisia. Tarkoituksena on selvittää, että voidaanko hälytin tunnistaa vaikka sen ääni olisi vaimentunut. Äänisignaalia analysoinnissa käytettiin apuna spektrogrammia. Spektrogrammista näkee riittävän tarkasti missä on suurin taajuus ja miten suuri se on muihin taajuuksiin verrattuna.

Testi osoitti, että laite itsessään ei vaimenna mitään taajuutta niin paljoa, että se häiritseisi signaalin havaitsemista. Jokaisella ajanhetkellä sinisignaali on selvästi havaittavissa. Havaittiin myös, että handsfree vaikuttaa signaalin jonkin verran mutta ei niin paljoa, että se häiritsisi signaalin havaitsemista. Handsfreellä ylimmät taajuudet, eli taajuudet lähellä 8000 hertsiä, näyttävät hieman vaimentuneen. Tämä ei häiritse mielenkiintoinen alue tämän työn kannalta on alle 5000 Hz.

Tämän testin perusteella voidaan sanoa, että laitteet itsessään eivät tee vaimenna mitään tiettyä taajuutta. Handsfree vaikuttaa ääneen hyvin vähän. Miltään taajuudelta vaimenneminen ei ole niin radikaalia, että tunnistaminen olisi mahdotonta.

5.3 Materiaalien aiheuttama vaimennus

Tätä diplomityötä varten testattiin miten erilaiset materiaalit vaikuttavat hälyttimen ääneen. Hälyttimien lisäksi testausta tehtiin kappaleessa 8.1 mainitulla sinisignaalin sweepillä ja valkoisella kohinalla. Testejä suoritettiin sekä kaiuttomassa huoneessa, että pihalla. Testeissä pyrittiin simuloimaan tilanteita missä hälyttimiä yleensä voitaisiin käyttää.

Edellisessä kappaleessa käytettiin sinisignaalia, jonka taajuus nousee yhdestä hertsistä 8000 hertsiin, testaamaan laitteen toimivuutta. Nyt sitä käytetään tarkastamaan miten eri materiaalit vaimentavat eri taajuuksia. Ääni nauhoitettiin laitteen omalla mikrofoniin ja handsfreen mikrofoniin. Testi suoritettiin yksinkertaisesti soittamalla sinisignaali ja sijoittamalla nauhoittava laite vaatteiden alle. Tämän jälkeen signaalista otettiin spektrogrammi josta näkee suurin piirtein mitkä taajuudet vaimenevat. Ensimmäisenä kokeiltiin farkkukankaiden läpi miten signaali vaimenee. Signaali vaimenee, mikäli vastaanottavan laitteen päälle laitetaan jotain. Tämä oli tietenkin odotettavissa. Taajuudet yli 7000 Hz näyttävät vaimenevan enemmän kuin alemmat. Signaali on kuitenkin edelleenkin hyvin havaittavissa ja sen tunnistaminen on mahdollista.

Seuraavassa testissä nauhoittava laite laitettiin laukun taskuun. Taskuja oli kaksi, uloin tasku missä vaimennusta on vähemmän ja sisin tasku, missä vaimennusta on eniten. Havaittiin, että signaali vaimenee enemmän taajuusalueen yläpäästä kuin alapäästä. Sisimmässä taskussa korkeammat taajuudet ovat jo selkeästi vaimentuneet verrattuna ei-vaimennettuun tilanteeseen. Sen sijaan taajuusalueen alapää on vaimentunut huomattavasti vähemmän vaikka silläkin alueella havaitaan vaimennukset vaikutukset. Erot näiden kahden taskun välillä on selvät ja laukun sisällä on äänen havaitseminen vaikeampaa. Laukun materiaali vaimentaa ääntä taajuuskaista yläpäästä enemmän kuin alapäästä.

5.4 Valkoinen kohina

Valkoista kohinaa nauhoitettiin samalla lailla kuin sinisignaalin sweep-testissä. Nauhoituksessa ääntä vaimennettiin farkkukankaalla ja laukulla. Laukun vaimennus on nauhoitettu eniten vaimentavassa taskussa eli sisimmässä taskussa. Testissä havaittiin laukun voimakas vaimentava vaikutus taajuusalueen yläpäähän. Samaan viittaa myös farkkujen aiheuttama vaimennus. Tulos on samankaltainen kuin signaalin sweep-testissä. Vaimennus kasvoi taajuusalueen alapäätä kohti sitä enemmän, mitä enemmän vaimentavaa materiaalia oli käytetty. Varsinkin liikkeessä on voi olla vaikeaa havaita signaalin ominaisuuksia, koska liikkeessä syntyy laukun sisällä myös ääntä. Farkkujen läpi ominaisuudet havaitaan helpommin.

5.5 Käden vaikutus

Käden vaimennusta ja mekaanisen puristuksen vaikutusta testattiin puristamalla erikoishälytin nyrkkiin. Tällainen tilanne, missä signaalia vaimentaa käyttäjän tai jonkun muun käsi, on hyvinkin mahdollinen ja on syytä testata miten nyrkkiin puristamien vaikuttaa ääneen. Testissä hälyttimen annettiin soida hetken aikaa, minkä jälkeen se puristettiin tiukasti nyrkkiin. Nauhoitus suoritettiin kaiuttomassa huoneessa. Tuloksiin vaikuttaa myös hälyttimeen kohdistuva mekaaninen puristus, mikä syntyy kun käsi puristetaan voimakkaasti nyrkkiin. Testissä havaittiin voimakasta vaimentumista laitetta puristettaessa. Kaikilla hälyttimillä havaittiin sama. Ryhmän 2 hälyttimille ominainen sahalaita kuvio oli kuitenkin vielä havaittavissa. Mekaanisesta puristuksesta johtuvia muutoksia äänessä ei havaittu. Voidaan olettaa, että mekaanisesta puristuksesta tulevat muutokset ovat erilaisia kuin vaimennuksesta johtuvat muutokset.

5.6 Heittotesti

Heittotestissä testattiin miten paljon hälyttimen ääni vaimenee kun se siirtyy nopeasti pois päin nauhoittavasta laitteesta. Testi suoritettiin pihalla erikoishälyttimelle. Testissä hälytin laitettiin soimaan ja siirrettiin 20 metrin päähän tasaisella vauhdilla. Testissä havaittiin signaalin voimakas vaimeneminen mitä kauemmas hälytin ja nauhoittava laite siirtyivät toisistaan. Hälyttimelle ominainen sahalaita on havaittavissa erittäin hyvin, kun etäisyys on alle 2 metriä ja tämän jälkeen ominaisuudet ovat koko ajan vaikeammin havaittavissa. Noin 10 metrin jälkeen äänen voimakkuus voi olla jo niin pieni, että laite ei välttämättä osaa tunnistaa sitä kovaksi ääneksi.

Kaikissa testeissä kyettiin havaitsemaan signaalin ominaisuuksia ainakin silmämääräisesti vaikka signaali oli vaimentunut voimakkaasti. Täydellisesti signaali ei näissä testeissä kadonnut. Ongelmaksi voi muodostua signaalin äänen voimakkuus, mikä luonnollisesti pienenee mitä enemmän signaalia vaimennetaan. Tällöin hälyttimen äänen voimakkuus voi olla niin pieni, että sitä ei havaita ollenkaan. On myös otettava huomioon mahdollinen taustamelu, jota näissä testeissä ei ollut kuin hyvin vähäisiä määriä. Vaimeneminen on siis otettava huomioon ja on syytä tarkkailla missä tilanteissa laitetta käytetään ja mihin laite sijoitetaan.

5.7 Piha

Erikoishälytin on käytössä enimmäkseen pihalla joten testejä suoritettiin niiden käyttäytymisestä pihalla. Hälyttimien äänet nauhoitettiin aukiolla alueella missä oli rakennuksia noin 40 metrin päässä. Nauhoitus tehtiin 1 metrin ja 10 metrin etäisyydellä hälyttimestä. Tarkoituksena on selvittää vaimeneeko ääni pihalla, onko pihalla enemmän taustamelua tai

muuta häiriötekijöitä ja voidaanko äänen ominaisuudet havaita edelleen. Testissä käytettiin palohälytin 1 ja erikoishälytintä. Testissä havaittiin, että mitä pihalla signaali aavistuksen puuroutuu. Tämä johtunee rakennuksista, mistä ääni kaikuu takaisin päin. Tämä vaikutus havaittiin vaikka rakennukset olivat melko kaukana testipaikasta. Tämä on sitä voimakkaampaa mitä kauempana hälytin on nauhoittavasta laitteesta. Toinen havaittava asia on se, että taustamelua löytyy jokaisesta kuvasta. Yleensä taustamelu on hälyttimen taajuusalueen alapuolella eikä vaikuta näin ollen signaalin ominaisuuksiin mutta on kuitenkin turhaa täytettä tunnistamisen kannalta. Erikoishälyttimessä havaitaan, että yhden metrin päässä signaalin sahalaita on selkeästi havaittavissa ja se nousee selvästi esille taustasta. Tämä on tärkeää, sillä erikoishälyttimet ovat yleensä lähellä laitetta.

Testin jälkeen voidaan sanoa, että pihalla mukaan tulee huomattavan paljon enemmän taustamelua kuin esimerkiksi kaiuttomassa huoneessa. Suurin osa tästä taustamelusta sijoittuu taajuudeltaan alle 1000 hertsiin eli siis hälytinäänien alapuolelle. Nämä voivat kuitenkin hävittää hälyttimen äänen allensa ja hälytyksen havaitseminen vaikeutuu.

5.8 Toimisto

Eräs toiminta-alue laitteelle on toimistotilat. Toimistotila on yleensä hieman kaikuvampi kuin normaali asuinhuoneisto ja toimiston huoneiden pinta-ala voi olla pienempi. Testissä hälyttimiä nauhoitettiin toimistossa 3 metrin etäisyydellä. Näin saadaan kuva siitä, miten hälytinäänäni käyttäytyy hieman kaikuvassa tilassa, joka kuitenkin on yksi mahdollisista toimintaympäristöistä. Testiin käytettiin palohälytin 1 ja erikoishälytintä. Testissä havaittiin kaikumisen vaikutus ääneen. Äänellisen jakson kaiku vaikuttaa äänettömään jaksoon. Kuvista voi vielä havaita äänettömän ja äänellisen jakson, mutta niiden laskeminen algoritmilla on huomattavan hankalaa. Palohälyttimen 1 kohdalla, missä näkyy samanlainen kaiun vaikutus äänettömään jaksoon. Erikoishälyttimellä havaittiin samankaltaisia ilmiöitä. Äänen taajuudet sekoittuvat kaiun vaikutuksesta eikä erikoishälyttimien sahalaita kuviota ole enää laskennallisesti helppo löytää. Sahalaidan havaitseminen on liki mahdotonta ihmissilmällekin. Tässä tapauksessa spektrogrammeissa näkyy vain suuren tehon alue sillä taajuusalueella missä on sahalaidan ylin ja alin taajuus. Voidaan sanoa, että signaalista ei enää kaikuvassa tilassa saada selville ominaisuuksia ainakaan niin luotettavasti kuin kaiuttomassa huoneessa. Toisaalta signaali itsessään ei katoa mihinkään ja se on havaittavissa oven sulkemisesta tai kaiusta huolimatta. Suurin tehon taajuusalue pystytään molemmissa ryhmissä havaitsemaan.

5.9 Erittäin kaikuva tila

Toimistossa suoritettuna testin lisäksi, suoritettiin testejä erittäin kaikuvassa tilassa. Testi suoritettiin pommisuojaissa. Kyseinen tila on tilavuudeltaan suuri, missä on etäisyydet seiniin suuret. Tila on louhittu maan alle, suoraan kallioon eli seinät heijastavat ääntä

erittäin hyvin. Toinen paikka oli putken mallinen ajokaista pommisuojaan. Seinät olivat kiveä, mutta tila oli pituuteensa nähden kapea. Pommisuojaassa testi suoritettiin 1 ja 10 metrin etäisyydellä ja ajoluiskassa, joka on putkimainen rakenne, 20 metrin etäisyydellä.

Palohälyttimen äänestä on hyvin vaikea löytää äänellinen jakso ja äänetön jakso, kunhan hälytin on noin metrin päässä nauhoittavasta laitteesta. Ihminen havaitsee vielä missä suurin piirtein jaksojen rajat menee. Tilanne kuitenkin ei ole niin selkeä kuin kaiuttomassa huoneessa. Tästä eteenpäin kaiun vaikutus on jo huomattavasti suurempaa. Kymmenen metrin päässä jaksot ovat jo lähestulkoon mahdottomia havaita. Signaalin teho on kuitenkin koko ajan havaittavissa ja sen alue on hyvin rajattu. Tämä tarkoittaa sitä, että ainakin yksi ominaisuus on mahdollista havaita eli signaalin taajuusalue. Putkimaisessa rakenteessa ääni pääsee kaikumaan enemmän ja äänestä ei enää havaita äänellisiä tai äänettä jaksot. Suurimman tehon omaava taajuusalue on havaittavissa.

Erikoishälyttimen kohdalla, pommisuojaassa tehdyssä testissä, on metrin päästä vielä havaittavissa sahalaita mutta etäisyyden kasvaessa ei tätä voida enää havaita mitenkään. Kaiku sotkee liiaksi hälytin ääntä ja signaali on muutenkin hankalampi kuin palohälyttimessä. Taajuusalue, millä on eniten tehoa, havaitaan sen sijaan helposti. Taajuuden alaraja näkyy hyvin selvästi kuvissa ja se on suhteellisen vakaa. Putkimainen rakenne on erikoishälyttimelle samankaltainen kuin palohälyttimelle. Äänestä on mahdoton havaita enää sahalaitaa mutta taajuusalue on vielä havaittavissa. Tässä se on noin 2200 hertsistä noin 5000 hertsiin. Hälyttimen äänessä olevasta sahalaidasta on kerrannaisia 5000 Hz yläpuolella ja nämä kerrannaiset vaikuttavat ylärajan havaitsemista.

Tämän testin perusteella voidaan sanoa, että kaikuvassa tilassa signaalin ominaisuuksien havaitseminen on hyvin vaikeaa, ellei jopa mahdotonta. Ainoa mikä näkyy selvästi, on taajuusalue, millä on suurin teho.

5.10 Hälytinlaitteiden testaus

Hälyttimen toimintaympäristö ei aina ole lämpötilaltaan tasainen. Tässä työssä testattiin, miten äärimmäisen kylmä lämpötila vaikuttaa hälyttimen ääneen olettaen, että kylmä ei muista syistä, esimerkiksi pariston loppumisen takia, lakkaa toimimasta. Hälyttimet olivat kylmässä (-23 °C) viisi vuorokautta. Tämän jälkeen hälyttimet jouduttiin siirtämään kaiuttomaan huoneeseen, minkä aikana ne lämpenivät jonkin verran. Tämä ei kuitenkaan poista mahdollisia pysyviä vahinkoja, mitä alhainen lämpötila on saattanut aiheuttaa. Palohälytin 1 toimi edelleen normaalilla tavalla. Havaittiin, että ominaisuudet eivät ole muuttuneet, esimerkiksi suurimman tehon omaavaa taajuus piirtyy spektrogrammeihin selkeästi ja sen taajuus ei ole muuttunut. Samaten äänelliset jaksot ovat pysyneet muuttumattomina. Erikoishälyttimen kohdalla havaittiin, että sahalaita on edelleenkin löydettävissä. Erikoishälyttimessä havaittiin pieniä ongelmia toiminnan

suhteen liittyen lähinnä paristoihin ja niiden kylmän kestävyysvarsinkin, jos paristo on jo käyttöikänsä loppupäässä. Tämä tarkoittaa sitä, että alhainen lämpötila saattaa vaikuttaa hälyttimien toimintaa vaikka niiden äänen ominaisuudet olisivatkin muuttumattomia.

6 YHTEENVETO

Palohälyttimien ominaisuuksia määrittävät standardit ovat melko tarkkoja ja niiden avulla voidaan tehdä tiettyjä oletuksia hälyttimien äänestä, sen voimakkuudesta ja kestosta. Kahta eri standardia käytetään Suomessa SFS-EN-14604 ja SFS-EN-54, jotka ovat suoraan Euroopassa käytössä olevista johdettu. Nämä määrittelevät nimenomaan palohälyttimien toimintaa ja nämä eivät koske esimerkiksi kaasuhälyttimiä tai erikoishälyttimiä.

Tämän työn tavoitteisiin kuuluu hälyttimien äänen tunnistaminen. Tähän liittyy tärkeänä osana signaalin ominaisuuksien tutkiminen ja ennen kaikkea löytää sellaiset piirteet, jotka poikkeavat joka hälyttimen kohdalla. Näin voidaan hälyttimet tunnistaa. Tätä varten jokaisen hälyttimen ääni analysointiin ja siitä etsittiin erilaisia ominaisuuksia. Ensimmäisessä vaiheessa palohälyttimestä, vesihälyttimestä ja muista ”piipittävästä” hälyttimistä otettiin vain yksi piirre, joka oli piipityksen taajuus. Tämä taajuus vaihteli jokaisella hälyttimellä ja näyttäisi pysyvän ajan suhteen vakiona. Toiseen vaiheeseen otettiin mukaan erikoishälyttimet, joiden ääni ei ole samanlainen ”piipittävä”. Näistä on hankalampaa tunnistaa ominaisuuksia, mutta jokaisella hälyttimellä on omalla taajuusalueella vaihteleva sahalaita aalto. Tämä aalto voidaan löytää ja ainakin sen taajuusalue selvittää. Mahdollista on myös selvittää sahalaidan jaksonaika, keskitaajuus ja huippujen taajuusero. ”Piipittävien” hälyttimien kohdalla löydettiin myös muita ominaisuuksia. Näistä voidaan mainita äänellisen jakson kesto, äänettömän jakson kesto ja äänien taajuus eli kuinka monta piipahdusta sekunnissa. Näistä voidaan luoda piirrevektoreita, joita voidaan käyttää tunnistamaan eri hälyttimiä.

Testejä tehtiin myös selvittääkseen miten eri materiaalit vaikuttavat hälytinten ääneen. Ääntä testattiin valkoisella kohinalla ja sinisignaallilla, jonka taajuutta nostettiin tasaisesti ylöspäin. Havaittiin, että taajuusalueen yläpäässä ääni vaimenee hiukan enemmän. Ongelmaksi voi muodostua se, että hälyttimen ääntä ei havaita ollenkaan vaimennuksen takia tai se peittyy taustameluun. Hälyttimiä testattiin myös pihalla, toimistossa ja erittäin kaikuvassa tilassa. Havaittiin, että kaikuminen voi hankaloittaa signaalin tunnistamista huomattavasti. Piirteet peittyvät herkästi kaiun alle. Ainoa ominaisuus mikä pysyi suhteellisen muuttumattomana, oli signaalin suurimman tehon omaava taajuusalue. Tämä taajuusalue on tärkein yksittäinen piirre, mihin signaalin tunnistaminen voi perustua.

Äkkinäisen muutoksen havaitseminen on helpohkoa, mikäli hälyttimen ääni on selvästi suurempi kuin ympäröivän tilan aiheuttama taustamelu. Yleensä näin on, sillä hälyttimet ovat kovaäänisiä. Muutoksen havaitsemiseen voidaan käyttää joko yhtä vakiinista kynnysarvoa. Tällöin oletetaan muutoksen tapahtuneen kun kynnysarvo ylittyy tai voidaan tehdä dynaaminen ja jakaa signaali ikkunoihin, joista lasketaan teho tai energia.

Tällöin voidaan vertailla signaalin nykyisen ikkunan tehoa tai energiaa edelliseen ikkunaan. Mukaan voidaan ottaa myös useampi edellinen ikkuna ja käyttää vertailuarvona näiden keskiarvoa.

LÄHTEET

- (Aumala, et. al. 1995) Aumala, O., Ihalainen, H., Jokinen, H., Kortelainen J. 1995. Mit-taussignaalien käsittely. 2. painos. Tampere. Pressus Oy. 317 p.
- (Bellanger 1989) Bellanger, M., 1989. Digital Processing of Signals, Theory and Prac-tice. 2. painos. Great Britain. John Wiley & Sons. 388 p.
- (Box et. al. 1997) Box G., Luceno, A. 1997. Statistical Control by Monitoring and Feedback Adjustment. United States of America. John Wiley & Sons, Inc. 327 p.
- (Brookes 2012) Brookes, M. Voicebox: Speech Processing Toolbox for Matlab. [www-dokumentti]. Imperial College, Department of Electrical & Electronic Engineer-ing. <<http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html>>. Luettu 17.08.2012.
- (DeFatta, et. al. 1988) DeFatta, D., Lucas, J., Hodgkiss W.. 1988. Digital Signal Proc-essing: A System Design Approach. Yhdysvallat. John Wiley & Sons, Inc. 661 p.
- (Dufaux 2001) Dufaux, A. 2001. Detection and Recognition of Impulsive Sound Sig-nals. Väitöskirja. Sveitsi. Neuchatelin yliopisto, mikroteknologian instituutti. 209 p.
- (Gray 1976) Gray, Jr, A., Markel, J. 1976. Distance Measures for Speech Processing. IEEE Transactions on Acoustics, Speech and Signal Processing. Vol. ASSP-35, No 5, pp 380-391.
- (Huttunen 2005) Huttunen, H. 2005. Signaalinkäsittelyn menetelmät. Tampereen teknil-linen yliopisto, Signaalinkäsittelyn laitos. Opintomoniste. 117 p.
- (Iser 2008) Iser, B., Schmidt, G., Minker, W. 2008. Bandwidth Extentsion of Speech Signals (Lecture Notes in Electrical Engineering). Germany. Springer. 182 p.
- (Jokinen 2003) Jokinen, H. 2003. Analysis of Non-stationary Measurement Signals: Aspects of Preprocessing and Spectrum Estimation. Väitöskirja. Tampere. Tam-pereen Teknillinen Yliopisto. Julkaisuja 441. 207p.
- (Kay 1988). Kay, S., M.. 1988. Modern Spectral Estimation, Theory & Application. Englewood Cliffs, New Jersey 07632. Prentice Hall. 543 p.

- (Kay 1993). Kay, S., M. 1993. Fundamentals of Statistical Signal Processing, Estimation Theory. Volume 1. Upper Saddle River, New Jersey, 07458. Prentice-Hall. 595 p.
- (Koutroumbas 2006) Koutroumbas, K., Theodoridis, S. 2006. Pattern Recognition. 3. painos. 525 B Street, Suite 1900, San Diego, CA 92101-4495, Yhdysvallat. Academic Press. 837 p.
- (Kuc, 1988) Kuc, R.. 1988. Introduction to digital signal processing. Yalen yliopisto, Yhdistyneet Kuningaskunnat. McGraw-Hill. 474 p.
- (Kwanerkaak et. al.1991) Kwanerkaak, H., Sivan, R. 1991. Modern Signals and Systems. Englewood Cliffs, New Jersey 07632. Prentice Hall. 791 p.
- (Lynn, et. al. 1994) Lynn, P., Fuerst, W. 1994. Introductory Digital Signal Processing with Computer Application. 2. painos. West Sussex, England. John Wiley & Sons. 479 p.
- (Mitra 2006) Mitra S., K. 2006. Digital Signal Processing, a Computer Based Approach. Department of Electrical and Computer Engineering, University of California, Santa Barbara. MCGraw-Hill. 972 p.
- (Oppenheim, et. al. 1975) Oppenheim, A., Schafer, R.. 1975. Digital Signal Processing. Englewood Cliffs, New Jersey, Yhdysvallat. Prentice-Hall, Inc. 585 p.
- (Oppenheim, et. al. 1989) Oppenheim, A., Schafer, R.. 1989. Discrete-Time Signal Processing. Englewood Cliffs, New Jersey, Yhdysvallat. Prentice-Hall, Inc. 879 p.
- (Oppenheim, et. al. 1997) Oppenheim, A., Willsky, A., Nawab, S.. 1997. Signals & Systems. 2. painos. Upper Saddle River, New Jersey, Yhdysvallat. Prentice-Hall, Inc. 957 p.
- (Phongsaphen 2006) Phongsaphen, J., Prapaisri, S. 2006. Least Squares and Discounted Least Squares in Autoregressive Process. Silpakorn University Internatiol Journal. Vol 6. Pp 122-135.
- (Porat 1994) Porat, Boaz. 1994. Digital Processing of Random Signals, theory and methods. Englewood Cliffs, New Jersey, Yhdysvallat. Prentice-Hall, Inc. 449 p.

- (Press, et. al. 1992). Press, W., Flannery, B., Teukolsky, S., Vetterling, W., Numerical Recipes in Fortran 77. 2. painos. University of Cambridge. Cambridge University Press. 1447 p.
- (Porat 1994) Porat, B. 1994. Digital Processing of Random Signals, Theory and Methods. Englewood Cliffs, New Jersey 07632. Prentice-Hall, Inc. 449 p.
- (Rossing, et. al. 2002) Rossing, Moore, Wheeler. 2002. The Science of Sound. 3. painos. Yhdysvallat. Pearson Education, Inc. 783 p.
- (Smith, 1997) Smith, S.. 1997. The Scientist and Engineer's Guide to Digital Signal Processing. California Technical Publishing. 640 p.
- (SFS-EN-14604, 2006) SFS-EN 14604. 2006. Palovaroittimet. Helsinki. Suomen Standardisoimisliitto. 115 p.
- (SFS-EN 54, 2011) SFS-EN 54. 2011. Palonhavaitsemis- ja palohälytysjärjestelmät. Helsinki. Suomen Standardisoimisliitto. 33 p.
- (Tohka 2012) Tohka, J. 2012. Johdatus Hahmontunnistukseen. Tampereen teknillinen yliopisto, Signaalinkäsittelyn laitos 2006-2012. Opintomoniste. 82 p.
- (Vaseghi 2000). Vaseghi S. 2000. Advanced Signal Processing and Noise Reduction. 2. painos. Brunel University, London, UK. John Wiley & Sons, Ltd. 514 p.
- (Wang 2003) Wang, A. 2003. Industrial-Strength Audio Search Algorithm. Proceedings of the Fourth International Conference on Music Information Retrieval. Baltimore, Yhdysvallat.
- (Wei 2000) Wei, B., Gibson J. Comparison of Distance Measures in Discrete Spectral Modeling. [www-dokumentti]. Department of Electrical Engineering, Southern Methodist University. Dallas, TX 75275.
<<http://spib.rice.edu/SPTM/DSP2000/submission/DSP/papers/paper125/paper125.pdf>> Luettu 27.8.2012.
- (Welch 1970) Welch, P.D. 1970. The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms. IEEE Trans. Audio Electroacoust., Vol. AU-15, pp 70-73.
- (Young 2008) Young, H., Freedman, R. Sears and Zemansky's University Physics with Modern Physics. 12. painos. San Francisco, Yhdysvallat. Pearson Education, Inc. 1551 p.

LIITE 1 TEHOSPEKTRIEN COSH- JA ITAKURA-SAITO- ALGORITMILLA LASKETUT ETÄISYYDET

Taulukossa on COSH- algoritmilla lasketut tehospektrien etäisyydet eri hälyttimien välillä. Tehospektrit ovat laskettu AR-kertoimien avulla. Merkintä "hf" tarkoittaa, että nauhoitus on tehty hansdfreen mikrofonilla.

COSH, AR Ord = 100	palohälytin 1	palohälytin 1, hf	palohälytin 2	palohälytin 2, hf	palohälytin 3	palohälytin 3, hf
palohälytin 1	0.00	0.49	24.91	46.15	39573.46	19090.64
palohälytin 1, hf	0.49	0.00	11.23	88.85	125176.61	57000.94
palohälytin 2	25.91	11.23	0.00	0.50	7693.57	5047.33
palohälytin 2, hf	46.15	88.85	0.50	0.00	3671.98	2477.42
palohälytin 3	39573.46	125176.61	7693.57	3671.98	0.00	0.57
palohälytin 3, hf	19090.64	57000.94	5047.33	2477.42	0.57	0.00
vesihälytin 2	410.73	1153.15	1934088.00	18.80	613.75	607.83
vesihälytin 2, hf	1604.30	4792.16	40.46	26.64	316.83	330.46
Erikoishälytin 6	327.07	619.71	596.56	127.56	128.67	139.06
Erikoishälytin 6, hf	283.23	339.61	488.82	126.42	297.11	186.94
Erikoishälytin 1	406.43	991.95	727.88	205.39	127.64	174.84
Erikoishälytin 1, hf	260.12	550.20	425.96	11630964.00	160.94	194.53

	vesihälytin 2	vesihälytin 2, hf	Erikoishälytin 6	Erikoishälytin 6, hf	Erikoishälytin 1	Erikoishälytin 1, hf
palohälytin 1	410.73	1604.30	327.07	283.23	406.43	260.12
palohälytin 1, hf	1153.15	4792.16	619.71	339.61	991.95	550.20
palohälytin 2	19.41	40.46	596.56	488.85	727.88	425.69
palohälytin 2, hf	18.80	26.64	127.56	126.42	205.39	116.10
palohälytin 3	613.75	316.83	128.67	297.11	127.67	160.94
palohälytin 3, hf	607.83	330.46	139.06	186.94	175.84	197.53
vesihälytin 2	0.00	0.17	127.41	275.73	132.01	179.53
vesihälytin 2, hf	0.17	0.00	112.08	157.61	108.59	106.04
Erikoishälytin 6	127.41	112.08	0.00	0.48	1.05	1.10
Erikoishälytin 6, hf	275.73	157.61	0.48	0.00	1.33	1.09
Erikoishälytin 1	132.01	108.59	1.05	1.33	0.00	0.31
Erikoishälytin 1, hf	179.53	106.04	1.10	1.09	0.31	0.00

Näiden arvojen laskentaan on käytetty puheen käsittelyyn soveltuvaa Matlabin työkalukokoelmaa nimeltään Voicebox ja sieltä funktiota distchar. Tämä laskee COSH-etäisyyden tehospektrien välille ja laskee tehospektrit AR-kertoimien avulla. Merkintä ”Ord” ilmoittaa, montako AR- kerrointa on käytetty. (Brookes 2012).

Taulukossa on Itakura- Saito- algoritmilla lasketut tehospektrien etäisyydet eri hälyttimien välillä. Tehospekt- rit ovat laskettu AR- kertoimien avulla. Merkintä "hf" tarkoittaa, että nauhoitus on tehty hansdfreen mikro- fonilla.

Itakura-Saito, AR, ord = 100	palohälytin 1	palohälytin 1, hf	palohälytin 2	palohälytin 2, hf	palohälytin 3	palohälytin 3, hf
palohälytin 1	0.00	0.40	8.26	15.47	588.46	1885.75
palohälytin 1, hf	0.58	0.00	3.63	18.41	149.86	535.64
palohälytin 2	41.56	18.84	0.00	0.42	270.22	924.39
palohälytin 2, hf	76.84	159.29	0.58	0.00	173.63	590.39
palohälytin 3	78558.47	250203.37	15116.93	7171.33	0.00	0.94
palohälytin 3, hf	36295.53	113466.23	9170.26	4364.36	0.21	0.00
vesihälytin 2	794.35	2300.89	35.09	33.93	22.32	91.87
vesihälytin 2, hf	3192.53	9580.69	78.20	50.88	21.67	88.00
Erikoishälytin 6	608.44	1198.47	1146.66	221.78	209.62	230.18
Erikoishälytin 6, hf	536.50	6483.72	936.26	227.20	521.83	301.48
Erikoishälytin 1	786.77	1959.36	1428.85	390.99	224.52	319.35
Erikoishälytin 1, hf	782.71	1065.31	820.89	205.77	294.54	360.85

	vesihälytin 2	vesihälytin 2, hf	Erikoishälytin 6	Erikoishälytin 6, hf	Erikoishälytin 1	Erikoishälytin 1, hf
palohälytin 1	27.11	16.07	45.70	29.96	26.09	37.53
palohälytin 1, hf	5.41	3.64	40.96	30.50	24.54	35.09
palohälytin 2	3.73	2.71	46.46	41.44	26.90	30.49
palohälytin 2, hf	3.66	2.41	33.63	25.63	19.79	26.42
palohälytin 3	1205.18	611.98	47.72	72.40	30.83	27.33
palohälytin 3, hf	1123.79	572.92	47.94	72.39	32.32	28.21
vesihälytin 2	0.00	0.17	46.33	54.86	27.02	26.28
vesihälytin 2, hf	0.16	0.00	36.89	44.83	22.68	18.35
Erikoishälytin 6	208.50	187.27	0.00	0.36	0.87	0.77
Erikoishälytin 6, hf	496.61	270.40	0.06	0.00	1.09	0.82
Erikoishälytin 1	237.00	194.50	1.23	1.57	0.00	0.28
Erikoishälytin 1, hf	332.77	193.72	1.43	1.36	0.34	0.00

Tehospektrien laskemiseen käytetty algoritmi laskee tehospektrin AR- kertoimien avulla. Merkintä ord ilmoittaa, montako AR-kerrointa on käytetty. Tässä tapauksessa on käytetty 100. Tähän on olemassa valmiita Matlab- algoritmeja sähköisessä muodossa. Näiden laskemiseen on käytetty puheen käsittelyyn soveltuvaa työkalukokoelmaa nimeltään Voicebox ja sen funktiota distisar. Tämä laskee tehospektrien etäisyyden käyttäen apunaan AR-kertoimia. (Brookes 2012).